The four height variables, boundary correlations, and dissipative defects in the Abelian sandpile model

M. Jeng*

Box 1654, Department of Physics, Southern Illinois University Edwardsville, Edwardsville, IL, 62025

We analyze the two-dimensional Abelian sandpile model, and demonstrate that the four height variables have different field identifications in the bulk, and along closed boundaries, but become identical, up to rescaling, along open boundaries. We consider two-point boundary correlations in detail, and discuss a number of complications that arise in the mapping from sandpile correlations to spanning tree correlations; the structure of our results suggests a conjecture that could greatly simplify future calculations. We find a number of three-point functions along closed boundaries, and propose closed boundary field identifications for the height variables. We analyze the effects of dissipative defect sites, at which the number of grains is not conserved, and show that dissipative defects along closed boundaries, and in the bulk, have no effect on any weakly allowed cluster variables, or on their correlations. Along open boundaries, we find a particularly simple field structure; we calculate all n-point correlations, for any combinations of height variables and dissipative defect sites, and find that all heights and defects are represented by the same field operator.

PACS numbers: 05.65.+b,45.70.-n

I. INTRODUCTION

The Abelian sandpile model (ASM), introduced by Bak, Tang, and Wiesenfeld, is the original prototype for self-organized criticality [1]. Systems with self-organized criticality are naturally driven to a critical point, and thus can potentially explain how power laws occur in nature without any fine-tuning of parameters. Since their introduction, sandpile models have been used to model an extraordinarily wide range of systems, from earthquakes [2] to river networks [3, 4]; see [5, 6] for reviews.

To be precise, we are considering the two-dimensional isotropic Abelian sandpile model. This is a very simple model; in fact, its simplicity is its strength, since otherwise it could not act as a model for such a diverse range of physical systems. The ASM is defined on a square lattice, where each site has a height variable (the number of grains of sand at that site) that can range from 1 to 4. At each time step, a grain of sand is added to a random site. Any site with more than four grains is unstable, and collapses, losing four grains, and sending one grain to each of its neighbors. Unstable sites are repeatedly collapsed until all sites are stable. Then, a new time step begins—a grain is added to a random site, and the process begins anew [1]. Initially, probabilities of configurations will depend on the initial conditions, but after a long period of time, the ASM develops a well-defined probability distribution of states, independent of the initial conditions [7]. Typically, the number of grains is conserved in each toppling, except for sites along open boundaries, where grains are lost with each toppling (i.e. fall off the edge). There must be at least one dissipative site—i.e. at least one site where the number of grains decreases upon toppling—or else the sandpile would eventually reach a state where topplings continued endlessly during a single time step.

Despite its simplicity, certain basic properties of the ASM remain unknown. For example, despite intensive work, the power law governing the sizes of avalanches in the ASM is still unknown—see [8] for a review. And while the height one variable is well understood, the roles played by the higher height variables (two, three, and four) are not. For example, no bulk two-point correlation functions of higher height variables are known.

It is known that the ASM is related to the set of spanning trees that can be drawn on the sandpile lattice, and that this relationship can be used to perform exact calculations of ASM probabilities [7, 9]. A spanning tree is a set of arrows drawn on the lattice, such that each site has exactly one arrow pointing from the site to a neighbor, and such that there are no closed loops of arrows. Following the path of arrows from any site will eventually lead off the edge of the sandpile (or, more generally, to a dissipative site, such as found on an open boundary)—the "site" off the edge of the sandpile is called the root. A number of relationships between the ASM and spanning tree states are known. For example, the number of recurrent states of the ASM (states that occur with nonzero probability after a long amount of time) is equal to the number of spanning trees that can be drawn on the sandpile lattice [9].

Spanning trees are, in turn, related to the $c = -2$ conformal field theory (CFT). The $c = -2$ CFT is the simplest known example of a logarithmic conformal field theory (LCFT), and is well understood [11–13].

A method introduced by Majumdar and Dhar exploits the mapping between ASM states and spanning tree states to obtain exact ASM probabilities [14]. It has long been known that the Majumdar-Dhar method can be used to find the two-point correlation function of the unit height variable, which decays as $1/r^4$ [14]. More recently, Mahieu and Ruelle used the Majumdar-

Dhar method to calculate correlation functions of a number of ASM height configurations, known as weakly allowed cluster variables [15]. They not only found that all the correlations decayed as $1/r^4$, but were able to use their correlations to identify the thirteen simplest weakly allowed cluster variables with operators in the LCFT. These variables were all identified with linear combinations of three LCFT field operators, all of which had scaling dimension two, but only one of which—the $\partial\theta\bar{\partial}\bar{\theta} + \bar{\partial}\theta\partial\bar{\theta}$ operator—was isotropic. In some ways, this suggested that the higher height variables should be identified with $\partial\theta\bar{\partial}\bar{\theta} + \bar{\partial}\theta\partial\bar{\theta}$; on the other hand, Mahieu and Ruelle pointed out that this appeared inconsistent with LCFT operator product expansions (OPE's).

Despite the power of these mappings, and of the Majumdar-Dhar method, fundamental questions about the ASM remain unanswered, because aspects of the mapping between the ASM and $c = -2$ LCFT are still unknown—for example, it is not known what field operators in the $c = -2$ LCFT represent the higher height variables of the ASM (or, indeed, whether such a representation even exists). A single site with height two, or any higher height variable, is not a weakly allowed cluster, and thus higher height probabilities and correlations cannot be calculated with the Majumdar-Dhar method. Priezzhev was able to extend the Majumdar-Dhar method to calculate the bulk probabilites for all higher height variables [10]. However, the bulk correlations of the higher height variables, which would be needed to obtain the field identifications of the higher height variables, remain unknown.

Ivashkevich calculated all two-point correlation functions of all height variables, along open and closed boundaries [16]. He found that all boundary correlations, between all height variables, decayed as $1/r^4$, and argued that this implied that all four height variables should be represented by the same field operator (up to rescaling). Dhar has argued that, based on clustering properties of correlation functions, the bulk correlations should be expected to factorize in a manner consistent with giving all four height variables the same field identification [17].

However, we argue here that the four height variables should in fact receive different field identifications, both along closed boundaries, and in the bulk, and propose field identifications along closed boundaries. Our conclusions are based on analysis of closed boundary three-point functions, and of dissipative defect sites, as well as a reanalysis of the methods and results of Ivashkevich. However, we show that along open boundaries all four height variables, as well as dissipative defect sites, are represented by the same operator, $\partial\theta\partial\bar{\theta}$, in the $c = -2$ LCFT. We demonstrate this by computing all n-point correlations of height variables and dissipative defects.

In sections II and III we briefly review the methods used by Majumdar, Dhar, and Priezzhev for studying the ASM. In section IV, we review Ivashkevich's calculations of the boundary height probabilities.

In section V, and appendices A and B, we discuss issues associated with boundary correlation functions. While Ivashkevich has already calculated the boundary two-point correlations [16], we show that he mischaracterized the mapping between ASM configurations and spanning tree configurations, and a correct characterization results in a number of complications, necessitating a reanalysis of the two-point correlations. The relationship between ASM states and spanning tree states is not what one might have initially expected; we also note that linear relationships between nonlocal spanning tree conditions and local spanning tree conditions for one-point probabilities do not carry over in a simple fashion for multipoint correlation functions. Both of these complications introduce what we call "anomalous graphs"—while these complications are important, because they are technical in nature, we delegate much of the discussion to the appendices. In section V we calculate the anomalous graphs, and conjecture that the anomalous graphs have no effect on the universal parts of any boundary correlation functions; while we have not been able to prove this conjecture, it holds true for all correlation functions that we have calculated.

In section VI, we look at correlation functions along closed boundaries. For two-point correlation functions, we find that while we disagree with Ivashkevich's relationship between ASM and spanning tree states, we agree with his final results. However, we argue that these final results are, in fact, not consistent with identifying all height variables with the same field operator. Next, we calculate all three-point functions along closed boundaries that involve at least one unit height variable, and use these to make field identifications along closed boundaries. Selected three-point functions appear in Eqs. (17-19), and we state the field identifications in Eqs. (20-22).

Next, in section VII, we introduce the concept of a dissipative defect site, and discuss its effect on the lattice Green functions for the open, closed, and bulk cases. In section VIII, we show that in the closed and bulk cases, dissipative defects have no effects on *any* weakly allowed cluster variables. This demonstrates that an analysis of weakly allowed cluster variables, such as that in [15], cannot provide a complete picture of the ASM. Our results imply, as a particular case, that dissipative defects in the closed and bulk cases have no effect on the unit height probability, or on correlations of unit heights. They do, however, have an effect on the higher height variables; we show this analytically for the closed case, in Eqs. (28-29), and have checked this numerically for the bulk case.

In sections IX, X, and XI, we compute all n-point correlation functions, for any number of height variables, and with any number of dissipative defects, along open boundaries. We find that there, all four height variables, and dissipative defects, are all represented by the same dimension two field, $\partial\theta\partial\bar{\theta}$. In fact, all local arrow diagrams along open boundaries are represented by $\partial\theta\partial\bar{\theta}$, up to multiplicative prefactors.

A short summary of these results can be found at [18].

II.   METHODS FOR ANALYSING THE ASM

At its core, the ASM is a tractable model because the sandpile model has an Abelian structure; the state of the sandpile does not depend on the order in which grains are added to the sites [7]. As a result of this Abelian structure, it can be shown that the states of the sandpile fall into two simple categories. Some of the 4^N states of the sandpile (where N is the number of sites) are transient, which means that they can occur early in the ASM's evolution, but occur with zero probability after an infinitely long time. The other states are recurrent, and all occur with equal probability after long times. So the probability for a property X to occur is nothing more than the fraction of recurrent states having property X.

To analyze the sandpile, it is convenient to allow more general toppling rules. We characterize the sandpile by a toppling matrix, $\Delta_{\vec{i}\,\vec{j}}$, where $\vec{i}$ and $\vec{j}$ are any lattice sites. $\vec{i}$ topples if its height is ever greater than $\Delta_{\vec{i}\,\vec{i}}$, at which point its height goes down by $\Delta_{\vec{i}\,\vec{i}}$, and the height of every other site $\vec{j}$ goes up by $-\Delta_{\vec{i}\,\vec{j}} \geq 0$ ($\Delta_{\vec{i}\,\vec{j}} = 0$ if $\vec{i}$ and $\vec{j}$ are not neighbors). The original ASM, described in the introduction, has $\Delta_{\vec{i}\,\vec{j}} = 4$ when $\vec{i} = \vec{j}$ (or $\Delta_{\vec{i}\,\vec{j}} = 3$ when $\vec{i} = \vec{j}$ is along a closed boundary), $\Delta_{\vec{i}\,\vec{j}} = -1$ when $\vec{i}$ and $\vec{j}$ are nearest neighbors, and $\Delta_{\vec{i}\,\vec{j}} = 0$ otherwise.

Dhar was able to show that the number of recurrent states, given very general restrictions on $\boldsymbol{\Delta}$, is equal to $\det(\boldsymbol{\Delta})$ [7]. However, $\det(\boldsymbol{\Delta})$ is also known to be equal to the number of spanning trees that can be drawn on the lattice [9]. In the spanning tree representation, $\Delta_{\vec{i}\,\vec{i}}$ indicates the number of neighbors that the arrow from $\vec{i}$ can point to, $\Delta_{\vec{i}\,\vec{j}} = -1$ if an arrow can point from $\vec{i}$ to $\vec{j}$, and $\Delta_{\vec{i}\,\vec{j}} = 0$ otherwise.

Certain height probabilities in the ASM can be equated with probabilities for spanning trees to have particular arrow configurations. Probabilities for *some* arrow configurations can be computed simply by modfiying the toppling matrix from $\boldsymbol{\Delta}$ to $\boldsymbol{\Delta}'$, in a way that enforces that arrow configuration. Then, the number of spanning trees with the configuration is $\det(\boldsymbol{\Delta}')$, and the probability of the configuration is $\det(\boldsymbol{\Delta}')/\det(\boldsymbol{\Delta})$. Defining $\mathbf{B} \equiv \boldsymbol{\Delta}' - \boldsymbol{\Delta}$, the probability becomes

$$\frac{\det(\boldsymbol{\Delta}')}{\det(\boldsymbol{\Delta})} = \det(\mathbb{I} + \mathbf{B}\mathbf{G}). \qquad (1)$$

$\mathbf{G} \equiv \boldsymbol{\Delta}^{-1}$ is the well-known lattice Green function [24] (see appendix C). If $\boldsymbol{\Delta}'$ only differs from $\boldsymbol{\Delta}$ in a finite number of entries, then $\mathbf{B}$ is finite-dimensional, and the probability can be easily computed.

Majumdar and Dhar used this method to find the probability for a site $\vec{i}$ to have unit height [14]. To do this, they defined a modified, or "cut" ASM, in which three of the four bonds connecting $\vec{i}$ to nearest neighbors are removed. When a bond is removed, the maximum height

of sites on each end is decreased by one; so the three sites adjacent to $\vec{i}$ get maximum heights of 3, and $\vec{i}$ gets a maximum height of 4-3=1. It is not difficult to show that recurrent states S (of the original ASM) where $\vec{i}$ has height one are in one-to-one correspondence with the recurrent states S' of the cut ASM. In this correspondence, we map from S to S' by lowering the heights of each of the three sites cut off from $\vec{i}$ by one. Letting $\vec{j}_1$, $\vec{j}_2$, and $\vec{j}_3$, be the three the neighbors that $\vec{i}$ has been cut off from, we have

$$\mathbf{B} = \begin{array}{cccc} \vec{i} & \vec{j}_1 & \vec{j}_2 & \vec{j}_3 \end{array} \\ \begin{pmatrix} -3 & 1 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix} \begin{array}{c} \vec{i} \\ \vec{j}_1 \\ \vec{j}_2 \\ \vec{j}_3 \end{array} \qquad (2)$$

Then the unit height probability is $\det(\mathbb{I} + \mathbf{B}\mathbf{G}) = 2(\pi - 2)/\pi^3$. This method was also used by Majumdar and Dhar to calculate the two-point correlation of the unit height variable [14].

Priezzhev extended the Majumdar-Dhar method to allow for the calculation of diagrams with closed loops. With the basic Majumdar-Dhar method, all off-diagonal entries of the toppling matrix are either 0 or -1. Priezzhev proved that if in $\boldsymbol{\Delta}'$ we set n off-diagonal entries of $\boldsymbol{\Delta}$ to $-\epsilon$, then

$$\lim_{\epsilon \to \infty} \frac{\det(\boldsymbol{\Delta}')}{\epsilon^n} \qquad (3)$$

is equal to the number of arrow configurations such that each of the n corresponding arrows is in a closed loop of arrows, where each closed loop contributes a factor of -1, and there are no closed loops other than those going through these n bonds.

Such configurations are not spanning trees; spanning trees cannot have any closed loops. However, Priezzhev found that to calculate certain spanning tree probabilities, he needed to calculate graphs that had closed loops (θ-graphs). We find this method useful for the calculation of certain closed boundary correlations.

III.   HEIGHT PROBABILITIES

Priezzhev determined a relationship between higher height probabilities and spanning tree states, which we review here [10].

Central to our analysis is the concept of forbidden subconfigurations (FSCs). A forbidden subconfiguration is a subset F of the lattice, such that for all $\vec{i} \in F$, $h_{\vec{i}} \leq c_{\vec{i}}(F)$, where $h_{\vec{i}}$ is the height of site $\vec{i}$, and $c_{\vec{i}}(F)$ is number of neighbors that $\vec{i}$ has in F. Majumdar and Dhar proved that a state of the ASM is recurrent if and only if it has no FSC's [7, 9].

The probability for a site $\vec{i}$ to have height two is more complicated than the height one probability [10]. In this case, changing the site height to one could either leave the ASM in an allowed (recurrent) state, or produce an FSC. The first case just gives the height one probability, which has already been calculated, so we consider the second case. Changing the height of $\vec{i}$ from two to one can produce multiple FSC's. Let F be the maximal forbidden subconfiguration (MFSC) produced by this change. (Because more than one FSC can be produced, the word "maximal" is necessary for complete precision, and for this mapping to work; Priezzhev simply referred to "the" FSC, but this does not introduce any errors in his analysis [10].) F must contain $\vec{i}$, and exactly one of the neighbors of $\vec{i}$, and be simply connected, but can otherwise have arbitrary shape. The states S of the original ASM where changing the height of $\vec{i}$ from two to one produces F as the MFSC are in one-to-one correspondence with states S' of a modified ASM. In the modified ASM, all the bonds bordering F are removed, except for one (arbitrarily chosen) bond of $\vec{i}$. In the correspondence, we map from S to S' by lowering heights of all sites that border F by the number of neighbors of F that they have been cut off from. (In this mapping, heights in F are unaffected.) With this mapping, the state S has no FSC's in the original ASM if and only if the state S' has no FSC's in the cut ASM. (Priezzhev's explanation used a slightly different, but equivalent, argument, based on the burning algorithm, a method for determining if a state is recurrent [7, 10]).

The site $\vec{i}$ is called a predecessor of the site $\vec{j}$ in the spanning tree if the path from $\vec{i}$ to the root goes through $\vec{j}$. We define $\mathrm{NNP}_{\vec{i}}$, as the number of nearest-neighbors of $\vec{i}$ that are predecessors of $\vec{i}$. Then, the correspondence above shows that the number of states of the modified ASM is equal to the number of spanning trees of the modified lattice, which is in turn equal to the number of spanning trees where F is the set of predecessors of $\vec{i}$. Summing over all possible sets F, we simply obtain the number of spanning trees where $\mathrm{NNP}_{\vec{i}} = 1$.

Similarly, it can be shown that the number of ASM states allowed when $\vec{i}$ has height h (or greater) but forbidden when $\vec{i}$ has height $h-1$ (or less), is equal to the number of spanning trees where $\mathrm{NNP}_{\vec{i}} = h-1$. Thus, the probability $P_{ASM}(h)$ for the site to have exactly height h in the ASM is

$$P_{ASM}(h) = \sum_{u=1}^{h} \frac{P_{SpTr}(u-1)}{m_{\vec{i}}+1-u} , \qquad (4)$$

where $P_{SpTr}(u-1)$ is the probability that a randomly chosen spanning tree will have $\mathrm{NNP}_{\vec{i}} = u-1$, and m_i is the maximum possible height of $\vec{i}$. ($m_{\vec{i}} = 4$ in the bulk, and along open boundaries, while $m_{\vec{i}} = 3$ along closed boundaries.) For more details, see [10].

This gives an exact representation of ASM height probabilities in terms of spanning tree probabilities. However,
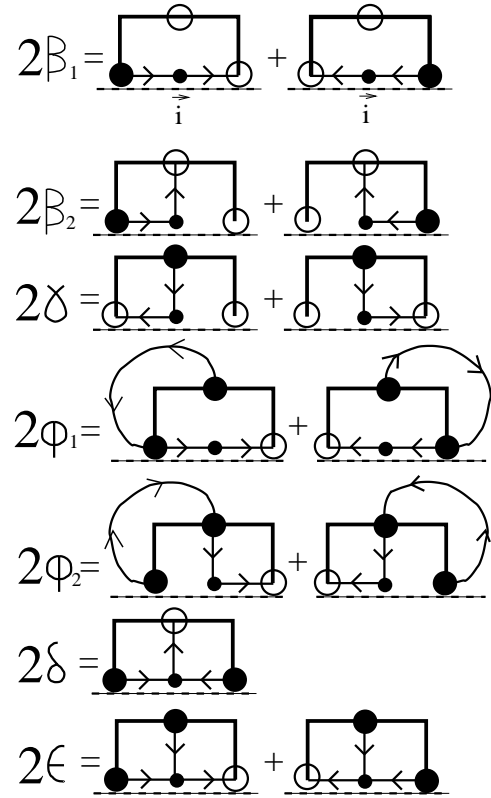

FIG. 1: Nonlocal arrow diagrams along closed boundaries.

these spanning tree probabilities are not easy to calculate. Spanning tree probabilities that correspond to local restrictions on the spanning tree can be calculated with the Majumdar-Dhar method. However, the statement that $\mathrm{NNP}_{\vec{i}} = u-1$ is a *nonlocal* restriction on the spanning tree (for $u > 1$). Priezzhev was able to calculate these nonlocal probabilities, but his calculations were complicated, and do not appear to be easily extensible to calculation of bulk correlations. However, this problem turns out to be more tractable along a boundary.

IV. BOUNDARY HEIGHT PROBABILITIES

For sites at the boundary, the relationship between height probabilities and NNP's still holds, and the NNP condition is still nonlocal. Nevertheless, Ivashkevich was able to show, through an ingenious transformation, that the ASM height probabilities are much easier to calculate along boundaries [16].

In figure 1 we list all possible nonlocal arrow configurations around a site $\vec{i}$ of a closed boundary. In each picture, the dashed line is the boundary, and the central site is $\vec{i}$. Large, solid, circles are predecessors of $\vec{i}$, while large, open, circles are not. We see explicitly that the predecessor relationships are nonlocal. β_1 and ϕ_1 differ only in whether the site above $\vec{i}$ leads to $\vec{i}$ by a chain of arrows—since the chain of arrows can go through sites
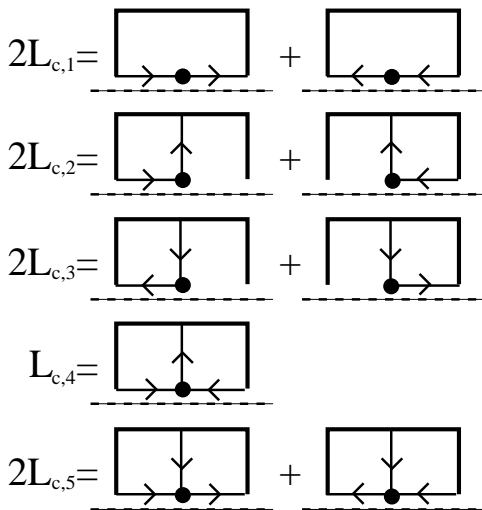
FIG. 2: Local arrow diagrams along closed boundaries.

FIG. 3: Anomalous graph of the first kind arising in the calculation of the two-point function.

distant from $\vec{i}$, this is a nonlocal distinction. If we can figure out the probabilities of all these diagrams, we can figure out the NNP probabilities (and thus the height probabilities). For example, the probability for $\text{NNP}_{\vec{i}} = 1$ is simply $2\beta_1 + 2\beta_2 + 2\gamma$, since these diagrams catalogue all the ways that $\vec{i}$ can have exactly one NNP.

These nonlocal diagrams are difficult to calculate. On the other hand, local restrictions are easily calculated with the Majumdar-Dhar method. All local arrow diagrams along closed boundaries are shown in figure 2. Note that these diagrams do not have solid or open circles, because predecessor relationships are not specified in local diagrams. Ivashkevich pointed out that the local arrow diagrams could be written as linear combinations of nonlocal arrow diagrams. For example, looking at figures 1 and 2, we see that $L_{c,1} = \phi_1 + \beta_1$. At first sight, there are more nonlocal arrow diagrams than local arrow diagrams, so such linear relationships would not appear to let us solve for the nonlocal arrow diagrams. However, Ivashkevich also pointed out that certain nonlocal arrow diagrams are equal in probability—for example, ϕ_1 and ϕ_2 are equal in probability, because we can make a one-to-one mapping from ϕ_1 to ϕ_2 by reversing all arrows in the long path of ϕ_1, and then switching the incoming arrow to $\vec{i}$. Similarly, $\beta_1 = \beta_2$. Then, we have as many nonlocal diagrams as local diagrams, and can solve for the nonlocal arrow diagrams. (In fact, along open boundaries, the number of local diagrams is one greater than the number of nonlocal arrow diagrams, so that the system is overconstrained, providing a check on the calculations.) Ivashkevich used this to calculate all height probabilities along open and closed boundaries. See [16] for the full list of linear relationships between local and nonlocal diagrams.
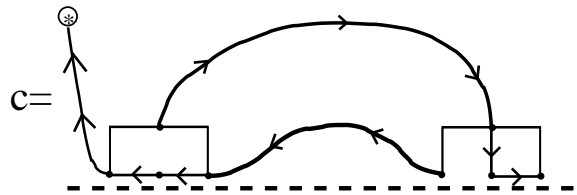
V. BOUNDARY TWO-POINT CORRELATIONS AND ANOMALOUS GRAPHS

The calculation of boundary correlations is much more difficult. We show in appendix A that Ivashkevich's calculation of the two-point functions was incorrect, and ignored complications that arise in the relationship between ASM height correlations and spanning tree correlations (although his end result turns out to be correct). In appendix B, we discuss further complications that arise in transforming from nonlocal spanning tree correlations to local spanning tree correlations. We summarize the results here, and analyze the resulting "anomalous graphs".

The first complication arises in the correspondence between ASM height probabilities and spanning tree probabilities. It would be natural to think that, analogously to Eq. (4), the ASM probability, $P_{ASM}(h_{\vec{i}}, h_{\vec{j}})$, for the sites $\vec{i}$ and $\vec{j}$ to have exactly heights $h_{\vec{i}}$ and $h_{\vec{j}}$ should be given by

$$P_{ASM}(h_{\vec{i}}, h_{\vec{j}}) \stackrel{?}{=} \sum_{u=1}^{h_{\vec{i}}} \sum_{v=1}^{h_{\vec{j}}} \frac{P_{SpTr}(u-1, v-1)}{(m_{\vec{i}}+1-u)(m_{\vec{j}}+1-v)} ,$$

(5)

where $P_{SpTr}(u-1, v-1)$ is the probability that in a spanning tree, $\text{NNP}_{\vec{i}} = u-1$ and $\text{NNP}_{\vec{j}} = v-1$. However, this turns out to not be quite the case. Eq. (5) is a natural guess, which we call a "naive" approach, but as shown in appendix A, the left and right sides of Eq. (5) differ by a subset of spanning trees that we call anomalous graphs of the first kind. (These graphs are not anomalous in any physical sense; we simply mean that they differ from what we would get, using a certain naive starting point.)

The set of anomalous graphs of the first kind, for the closed case, is shown in figure 3. (We represent the root with a star.) In the graph, $\text{NNP}_{\vec{i}} = \text{NNP}_{\vec{j}} = 1$, so this graph appears in the right-hand side of Eq. (5) for $h_{\vec{i}} = h_{\vec{j}} = 2$. However, we show in appendix A that this graph does not contribute to the 2-2 (height two-height two) correlation, but instead contributes to the 2-3 and 3-2 correlations, and gets subtracted from the 3-3 correlation.

Second, leaving aside for now the anomalous graphs of the first kind, we need to calculate correlations of nonlocal arrow diagrams. It would be convenient if we could use the linear relationships relating nonlocal arrow diagrams to local arrow diagrams found for one-point func-
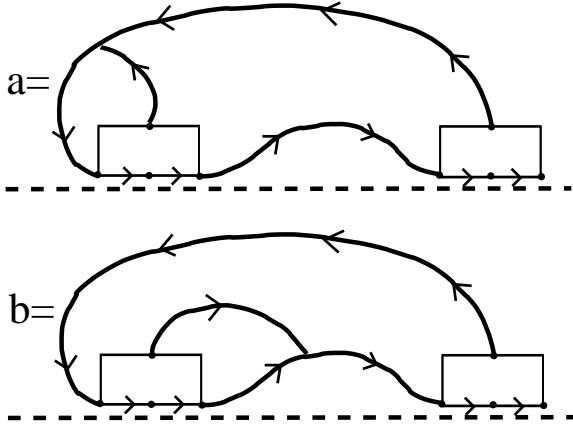
FIG. 4: Anomalous graphs of the second kind arising in the calculation of the two-point functions.
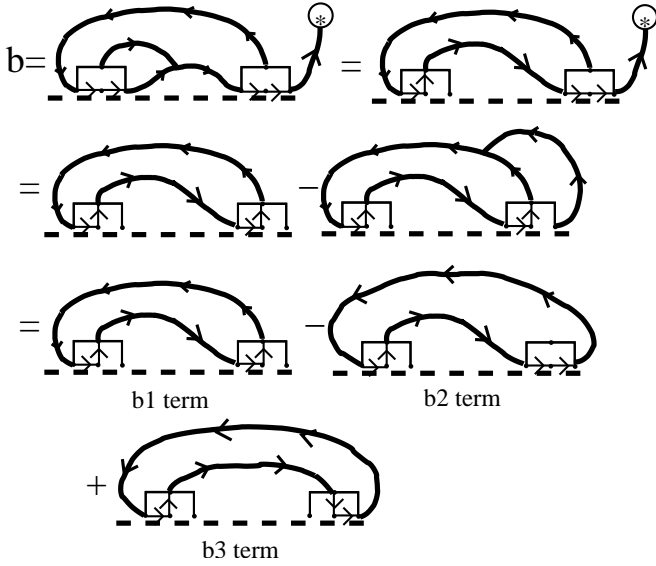

FIG. 5: Anomalous graph b as a linear combination of closed loop diagrams.

tions (section IV), and use them *independently* at $\vec{i}$ and at $\vec{j}$ for two-point functions. We again call this approach "naive," and again, this approach does not quite work. The problem arises because for one-point functions, we treated ϕ_1 and ϕ_2 as equivalent, based on a one-to-one correspondence in which a long path was reversed. In a correlation function of nonlocal arrow diagrams, the long path from a ϕ_1 at $\vec{i}$ may go through arrow constraints near $\vec{j}$, which are not free to be reversed. We discuss this problem in detail in appendix B. Consideration of this problem shows that, relative to the naive approach, our results are changed by graphs a and b, shown in figure 4. We call these anomalous graphs of the second kind.

The anomalous graphs a, b, and c can be calculated with the extension of the method of Priezzhev, discussed in section II [10]. We discuss only the calculation of the b term; the analysis of the other terms is similar.

b represents a subset of spanning trees, and thus cannot have any closed loops. However, it comes "very close" to having a closed loop that includes the distant sites $\vec{i}$ and $\vec{j}$, and we see in figure 5 that b can be written as a sum of closed loop diagrams.

Priezzhev's method allows us to calculate the closed loop diagrams. We represent an arrow whose weight in $\mathbf{\Delta}$ is set to $-\epsilon$ ($\epsilon \to \infty$) with a wavy bond line. As discussed in section II, these bonds must be part of a closed loop, and we get a factor of -1 for every closed loop. This gives the relations in figure 6. Taking the difference of the two graphs in figure 6 then gives the value of a closed loop diagrams that goes through both $\vec{i}$ and $\vec{j}$. Using this method, we find the number of diagrams $b1$, $b2$, and $b3$ (labeled in figure 5), as ratios of N, the total number of spanning trees:

$$\frac{N_{b1}}{N} = \frac{(3\pi-8)^2(\pi g(x)-1)}{4\pi^4 x^2} + \frac{(-128+48\pi+\pi^2)+(256-192\pi+30\pi^2)\pi g(x)}{16\pi x^4} + \mathcal{O}(\frac{1}{x^6}) \tag{6}$$

$$\frac{N_{b2}}{N} = \frac{(3\pi-8)^2(\pi g(x)-1)}{2\pi^4 x^2} + \frac{(3\pi-8)(2\pi g(x)-1)}{4\pi^3 x^3} + \frac{(3\pi-8)((4-\pi)+(3\pi-8)\pi g(x))}{2\pi^4 x^4} + \mathcal{O}(\frac{1}{x^5}) \tag{7}$$

$$\frac{N_{b3}}{N} = \frac{(3\pi-8)^2(\pi g(x)-1)}{4\pi^4 x^2} + \frac{(3\pi-8)(2\pi g(x)-1)}{4\pi^3 x^3} + \frac{(-128+48\pi-\pi^2)+(256-192\pi+42\pi^2)\pi g(x)}{16\pi x^4} + \mathcal{O}(\frac{1}{x^5}) \tag{8}$$

x is the separation between $\vec{i}$ and $\vec{j}$ along the defect. $g(x)$ is the Green function between $\vec{i}$ and $\vec{j}$, and diverges as $\ln(L)$, where L is the system size (it also diverges as $\ln(x)$). The restriction that spanning trees should have no closed loops greatly limits the number of possible spanning trees, when the outlets to the root (open boundaries) are very far away. So diagrams such as $b1$, $b2$, and $b3$, that allow a closed loop, are much more numerous than diagrams of spanning trees.

However, to find b, we take the linear combination,
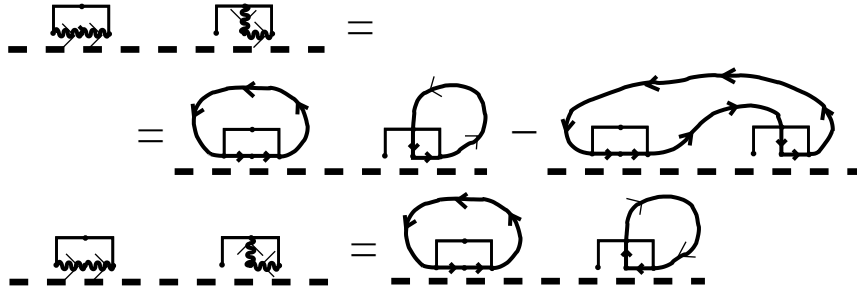
FIG. 6: Use of -ϵ weight bonds to evaluate closed loop diagrams.

$(N_{b1} - N_{b2} + N_{b3})/N$, and the $\ln(L)$ divergences cancel (this provides a check on our calculations). a and c can be found similarly. We find

$$a = \frac{3\pi - 8}{2\pi^3 x^4} + \mathcal{O}(\frac{1}{x^5}) \tag{9}$$

$$b = \frac{3\pi - 8}{2\pi^3 x^4} + \mathcal{O}(\frac{1}{x^5}) \tag{10}$$

$$c = \mathcal{O}(\frac{1}{x^6}) \tag{11}$$

a and b are both of order $1/x^4$. The two-point correlation functions turn out to decay as $1/x^4$, so the anomalous graphs could, in principle, affect the universal parts of the correlation functions. However, the anomalous graphs of the second kind come in the combination $(a-b)$ (see Eqs. (B5-B7)). So their total contributions to the two-point correlations are $\mathcal{O}(1/x^5)$, and can be dropped.

The end result is rather surprising. A "naive" approach might simply apply the relationship between ASM states and NNP conditions found for the one-point functions, independently at $\vec{i}$ and $\vec{j}$ (i.e. extend Eq. (4) to Eq. (5)), and then apply the relationships between non-local arrow diagrams and local arrow diagrams found for the one-point functions, independently at $\vec{i}$ and $\vec{j}$. Neither of these steps is correct, and a correct analysis produces correction terms (the anomalous graphs) to this naive approach. But, somehow, the anomalous graphs, while nonzero, produce no correction to the leading-order, universal results at any stage of the computation; the naive approach gives the answers. In fact, we find in the following sections that the naive approach again gives correct results for all three-point closed boundary correlations that we have calculated, and for all open boundary correlations. This leads us to conjecture that the naive approach always produces correct universal results, for all correlations. If this conjecture were proven true, it would greatly simplify further calculations—for example, the anomalous diagrams have prevented us from calculating the 2-2-2 correlation along closed boundaries.

VI. TWO- AND THREE-POINT CLOSED BOUNDARY CORRELATION FUNCTIONS

We define, for all correlation functions along closed boundaries,

$$\begin{aligned} f_{\rm c}(a_1, a_2, \ldots, a_n) &= \\ &\langle (\delta_{h_{x_1}, a_1} - p_{a_1, {\rm c}}) \ldots (\delta_{h_{x_n}, a_n} - p_{a_n, {\rm c}}) \rangle_{\rm c} , \end{aligned} \tag{12}$$

In this correlation function, the height h_{x_u} at the boundary site x_u is required to be a_u. We have subtracted off the constant boundary probabilities, $p_{a_u, {\rm c}}$, which were found in [16], as described in section IV. The subscript "c" stands for "closed." As already noted, despite errors in the setup in [16], the results of [16] are nevertheless correct, where it was found that

$$f_{\rm c}(1,1) = \left(-\frac{9}{\pi^2} + \frac{48}{\pi^3} - \frac{64}{\pi^4} \right) \frac{1}{(x_1 - x_2)^4} + \ldots \tag{13}$$

$$f_{\rm c}(1,2) = \left(\frac{12}{\pi^2} - \frac{68}{\pi^3} + \frac{96}{\pi^4} \right) \frac{1}{(x_1 - x_2)^4} + \ldots \tag{14}$$

$$f_{\rm c}(2,2) = \left(-\frac{61}{4\pi^2} + \frac{96}{\pi^3} - \frac{144}{\pi^4} \right) \frac{1}{(x_1 - x_2)^4} + \ldots \tag{15}$$

The correlation functions involving the height three variables were also calculated, but we do not write them here, as they are determined by the requirement that all height probabilities must sum to one at every site. (There is a misprint in the result for $f_{\rm c}(3,3)$ in [16].)

Ivashkevich argued that the fact that all two-point correlations decay as $1/(x_1 - x_2)^4$ indicates that all three height variables are represented by the same field operator. However, if all height variables were represented by the same operator, we would expect the two-point functions to factorize, as

$$f_{\rm c}(u,v) = -\frac{K_u K_v}{(x_1 - x_2)^4} \quad ; \quad u,v \in \{1,2,3\} , \tag{16}$$

for some constants K_u. However, the results in Eqs. (13-15) do not factorize in this manner. Dhar argued that

we should expect this factorization for bulk correlations, based on the "clustering properties of correlation functions," but we see that this factorization already fails along closed boundaries [17]. (We will see later that the open boundary correlations do, however, factorize in this manner, for all n-point correlations.)

To clarify the field identifications, we have calculated all three-point functions along closed boundaries, where at least one of the heights is the unit height. Some of our results are

$$f_c(1,1,1) = \frac{2(3\pi-8)^3}{\pi^6(x_1-x_2)^2(x_1-x_3)^2(x_2-x_3)^2} + \ldots \tag{17}$$

$$f_c(1,1,2) = -\frac{8(\pi-3)(3\pi-8)^2}{\pi^6(x_1-x_2)^2(x_1-x_3)^2(x_2-x_3)^2}$$
$$-\frac{(3\pi-8)^2}{\pi^5(x_1-x_3)^3(x_2-x_3)^3} + \ldots \tag{18}$$

$$f_c(1,2,2) = -\frac{4(3\pi-8)(-5\pi^2+39\pi-72)}{\pi^6(x_1-x_2)^2(x_1-x_3)^2(x_2-x_3)^2}$$
$$+\frac{(3\pi-8)(24-7\pi)}{2\pi^5(x_1-x_2)^3(x_1-x_3)^3} + \ldots \tag{19}$$

Other three-point correlation functions, calculated with the same methods, are listed in appendix D. They are all consistent with the requirement that the three height probabilities must sum to one at any site, and permutation symmetry, thus providing a check on our calculations.

We again get a number of anomalous graphs (relative to a naive approach), and as stated in the previous section, again find that all anomalous graphs cancel in the universal, leading-order terms of the correlation function.

These correlation functions are consistent with identifying the height variables with the following field operators in the $c=-2$ CFT:

$$\text{Height one} \;:\; -\frac{2(3\pi-8)}{\pi^2}\left(\partial\theta\partial\bar{\theta}\right) \tag{20}$$

$$\text{Height two} \;:\; \frac{6(\pi-4)}{\pi^2}\left(\partial\theta\partial\bar{\theta}\right) + \frac{1}{2\pi}\theta\partial^2\bar{\theta} \tag{21}$$

$$\text{Height three} \;:\; \frac{8}{\pi^2}\left(\partial\theta\partial\bar{\theta}\right) - \frac{1}{2\pi}\theta\partial^2\bar{\theta} \tag{22}$$

The representation of the $c=-2$ CFT used here is described briefly in appendix E. Note that the boundary correlations in Eqs. (17-19) are the same as the *bulk* correlations of Eqs. (20-22), and that while the $c=-2$ CFT contains holomorphic and antiholomorphic fields (the ∂ and $\bar{\partial}$ of Eq. (E1)), the fields in Eqs. (20-22) contain only holomorphic fields. This is consistent with boundary CFT. While fields in the bulk generally have holomorphic and antiholomorphic parts, near a boundary the antiholomorphic pieces behave, in all correlation functions, like holomorphic pieces at mirror positions across the boundary [19].

It is also consistent to make the substitution $\theta \to \bar{\theta}$, $\bar{\theta} \to -\theta$ in these field identifications, as the $c=-2$ LCFT is symmetric under this transformation (see Eq. (E1)).

The fact that the field identifications for the height variables differ along a closed boundary proves that they must also differ in the bulk. This is because in a CFT, boundary operators are derived from operator product expansions (OPE's) of bulk operators [19]. Furthermore, in appendix F we present a simple argument, based on general CFT principles, and not on any detailed calculations, that the height variables must have different field identifications in the bulk.

We have not been able to calculate three-point correlation functions that have no unit height variables. The basic problem is with the anomalous diagrams that arise when we convert from nonlocal arrow diagrams to local arrow diagrams (as in appendix B). The trick shown in figure 6, for evaluating the resultant closed loop diagrams, does not work for these three-point functions. We note that if we use the conjecture proposed in section V (i.e., ignore the anomalous graphs), we obtain

$$f_c(2,2,2) = -\frac{(24-5\pi)(-576+384\pi-61\pi^2)}{4\pi^6(x_1-x_2)^2(x_1-x_3)^2(x_2-x_3)^2} + \ldots, \tag{23}$$

(and other three-point functions consistent with the requirement that all three height probabilities must sum to one at any site). This correlation function is consistent with the field identification in Eq. (21), providing support for our conjecture.

VII. DISSIPATIVE DEFECT SITES, GENERAL

We now consider the effects of dissipative defects on the ASM. Generally, at sites in the bulk, or along closed boundaries, the number of grains is conserved at each toppling. Usually, it is only at open boundaries that the number of grains is not conserved; there, of the four grains toppled, three grains are sent to neighbors, while the remaining grain goes off the edge of the sandpile, to the root.

Some dissipation (i.e. sites where topplings remove grains from the sandpile) is necessary for the sandpile model to be well defined, since otherwise we would end up with states where the topplings never terminated. Nevertheless, dissipation often plays a minor role in analysis of the sandpile, because properties are often studied in the bulk of the ASM, with the dissipative sites along the open boundaries infinitely far away.

Some previous studies have investigated the effect of adding dissipation throughout the bulk of the ASM. Instead of having bulk sites topple when their height is greater than 4, they topple when their height is greater than $4+k$ ($k>0$). Then, one grain is sent to each of the four neighbors, and k grains are lost to the root. It has been shown, both numerically and analytically,

that if this is done at all sites, the ASM is taken off the critical point, and the power law correlations are destroyed [20–22]. This happens even when k is infinitesimal. (Although this modification to the ASM has its most obvious interpretation for integer k, the theory can be given a sensible interpretation for any rational value of k. See [20] for details.) More recently, Mahieu and Ruelle have demonstrated the precise manner in which dissipation throughout the bulk takes the ASM off the critical point. They found that the dissipation has exactly the same effect on correlation functions of weakly allowed cluster variables, as adding the integral of the dimension 0 variable, $\theta\bar{\theta}$, to the $c = -2$ CFT [15]. Adding dissipation along a line has been shown to split the ASM into two separate half-planes, each with open boundary conditions [23].

Here, we consider the effect of adding dissipation at only a single defect site. Then, the methods of Majumdar and Dhar still work, but we need to use a modified lattice Green function. If k grains of sand are dissipated at the lattice position $\vec{d}$, then we call k the "strength" of the defect. The toppling matrix is then changed from the defect-free toppling matrix, $\mathbf{\Delta}_0$, to

$$\Delta_{\vec{i},\vec{j}} = \Delta_{0;\vec{i},\vec{j}} + k\delta_{\vec{i},\vec{d}}\,\delta_{\vec{j},\vec{d}} \tag{24}$$

The Green function is simply the inverse of the toppling matrix, and is changed from the defect-free Green function, G_0 (described in appendix C), to

$$G(\vec{i},\vec{j}) = G_0(\vec{i},\vec{j}) - \frac{k}{1 + kG_0(\vec{d},\vec{d})}G_0(\vec{i},\vec{d})G_0(\vec{d},\vec{j}) \tag{25}$$

This holds for any value of k, and regardless of the location of the defect. Nevertheless, the defect behaves very differently in the open case, and in the closed/bulk cases. This is because the Green function between nearby lattice sites is $\mathcal{O}(1)$ near an open boundary, but $\mathcal{O}(\ln L)$ near a closed boundary, or in the bulk [24, 25]. L is the system size, or more generally, is of the same order-of-magnitude as the distance to the nearest open boundary. This divergence in the Green function as $L \to \infty$ for the closed and bulk cases is usually not an issue, since in most cases, we are concerned with differences in Green functions. However, here the divergence of all the $\mathbf{G}_0$ terms makes Eq. (25) unwieldy, although technically correct. (Eq. (25) can be used in the open case without modification.) We work in a limit where the distances between $\vec{i}$, $\vec{j}$, and $\vec{d}$, while possibly large, are all much less than L. In this limit, dropping terms of order $1/(\ln L)$, Eq. (25) becomes

$$G(\vec{i},\vec{j}) = G_0(\vec{i},\vec{j}) - G_0(\vec{i},\vec{d}) - G_0(\vec{d},\vec{j}) + G_0(\vec{d},\vec{d}) \tag{26}$$

Note that Eq. (26) is independent of k. This makes sense, since in the bulk, or along a closed boundary, spanning trees have to travel far to reach the root. But with the defect given by Eq. (25), k bonds are added from the defect $\vec{d}$ to the root. Adding a dissipative defect provides such an "easy" way to reach the root, that with high probability (probability one as $L \to \infty$), all nearby points will be predecessors of the dissipative defect, regardless of the value of k. The set of spanning trees will thus be the same, in the $L \to \infty$ limit, for any k. Note also that the Green function in Eq. (26) no longer diverges as $L \to \infty$, which is appropriate, as we are no longer $\mathcal{O}(L)$ from any dissipative sites.

VIII. DISSIPATIVE DEFECT SITES, CLOSED AND BULK CASES

Surprisingly, it turns out that a dissipative defect, either in the bulk, or on or near a closed boundary, has no effect on any weakly allowed cluster variables in the ASM. Weakly allowed cluster variables are height configurations that result in a subconfiguration that contains an FSC if any height in the configuration is reduced by one [26]. Examples of weakly allowed cluster variables are a single height one variable, or a height one adjacent to a height two. Such variables can be calculated with the Majumdar-Dhar method by the removal of a set of bonds in the ASM/spanning tree. We note that correlations of weakly allowed cluster variables (such as all correlations of the unit height variable) are also weakly allowed cluster variables.

Probabilities of weakly allowed cluster variables can be calculated as $\det(\mathbb{I} + \mathbf{BG})$, as in section II. To analyze the effects of the defect, we want to consider the effect of modifying the Green function from the defect-free Green function $\mathbf{G}_0$, to the Green function in Eq. (26), for a fixed matrix $\mathbf{B}$ (i.e. for a specific weakly allowed cluster variable).

In general, for local arrow restrictions, each row of $\mathbf{B}$ must sum to zero, because if the restrictions on the spanning trees prevent an arrow from $\vec{i}$ to $\vec{j}$, then $B_{\vec{i},\vec{i}}$ goes down by 1, while $B_{\vec{i},\vec{j}}$ goes up by 1. (For example, for the height one variable, the matrix $\mathbf{B}$ in Eq. (2) arises from the restriction that no arrows can point from $\vec{i}$ to $\vec{j}_1$, $\vec{j}_2$, or $\vec{j}_3$, nor from $\vec{j}_1$, $\vec{j}_2$, or $\vec{j}_3$ to $\vec{i}$.) For the weakly allowed cluster variables, $\mathbf{B}$ is symmetric, since if the arrow from $\vec{i}$ to $\vec{j}$ is forbidden, then so is the arrow from $\vec{j}$ to $\vec{i}$. So every column of $\mathbf{B}$ also sums to zero.

Since every row of $\mathbf{B}$ sums to zero, the parts of $\mathbf{G}$ that are independent of the row index of $\mathbf{G}$ make no contribution to $\mathbf{BG}$, and thus no contribution to the probability $\det(\mathbb{I} + \mathbf{BG})$. And $\det(\mathbb{I} + \mathbf{BG}) = \det(\mathbb{I} + \mathbf{GB})$, so since every column of $\mathbf{B}$ sums to zero, the parts of $\mathbf{G}$ that are independent of the column index also make no contribution to the probability. The last three terms of Eq. (26) all depend either only on the row index, or only on the column index. So a dissipative defect has no effect on any weakly allowed cluster probabilities (either on one-point probabilities or on correlations).

As a special case, this means that the unit height probability, and its correlations, are unaffected by closed or bulk dissipative defects. However, the higher height variables are affected. Using the Green function in Eq. (26), and the methods described in section IV, we find that along a closed boundary, with a defect at the origin, we have the following height probabilites at x_1:

$$f_c(1) = 0 \qquad (27)$$

$$f_c(2) = -\frac{1}{2\pi x_1^2} + \ldots \qquad (28)$$

$$f_c(3) = +\frac{1}{2\pi x_1^2} + \ldots \qquad (29)$$

We have numerically confirmed these results. These results provide further evidence that the height two and three variables have different field identifications along closed boundaries.

Since the height two and three variables have dimension two, this indicates that a dissipative defect along a closed boundary is a dimension zero operator. Consistent with this, uniform dissipation in the bulk has been identified with the integral of the dimension zero operator $\theta\bar{\theta}$ [15]. However, the correlation of $\theta\bar{\theta}$ with the height two and three operators in Eqs. (21-22) does not produce the correlations in Eqs. (27-29); this situation requires further analysis.

In the bulk, we would also expect that the higher height probabilites would be affected by a defect site, and have confirmed this with numerical simulations, although have not proven this analytically.

The fact that weakly allowed cluster variables have no correlations with bulk or closed defects provides compelling evidence that weakly allowed cluster variables do not provide a complete picture of the sandpile model. This has particular bearing on the analysis of Mahieu and Ruelle [15]. They studied specific bulk correlations of the simplest weakly allowed cluster variables, and developed a complete field picture for these variables. They found that (at the critical point) these variables are all linear combinations of three dimension two variables, $\partial\theta\bar{\partial}\bar{\theta} + \bar{\partial}\theta\partial\bar{\theta}$, $\partial\theta\partial\bar{\theta}$, and $\bar{\partial}\theta\bar{\partial}\bar{\theta}$, strongly indicating that all weakly allowed cluster variables are linear combinations of these three fields. However, this analysis left the status of the height two variable unresolved. Mahieu and Ruelle pointed out that since the height two variable appears in a number of the weakly allowed cluster variables, it might be expected that the height two variable would also be a linear combination of these three fields, or more specifically, proportional to the sole rotationally invariant field, $\partial\theta\bar{\partial}\bar{\theta} + \bar{\partial}\theta\partial\bar{\theta}$ [15]. But they also noted that such an identification appeared inconsistent with the fusion rules of the $c = -2$ CFT, which would indicate a different field identification. The analysis here points strongly to the latter conclusion, although the specific field identification in the bulk remains unresolved.

IX. ALL n-POINT CORRELATIONS ALONG OPEN BOUNDARIES, PART I

We have calculated all n-point correlations of all four height variables, along open boundaries, in the presence of an arbitrary number of dissipative defects. We begin by discussing why this case is so tractable (in contrast to the closed case, where we have been unable to calculate the three-point function of the height two variable).

The heights of the correlation function are placed at $x_1, x_2, \cdots, x_n$, and defining $x_{ab} \equiv x_a - x_b \equiv c_{ab}x$, we work in the limit $x \to \infty$, where the c_{ab}'s are kept constant.

As discussed in section V, and appendices A and B, a number of anomalous terms arise in the computation of correlation functions. While the discussion in these sections focused on closed boundary correlations, similar anomalous graphs arise in open boundary correlations. However, it turns out that these anomalous graphs produce no contributions to the universal parts of any correlation functions, greatly simplifying matters. We prove this claim in this section, and in the next section look at the actual calculation of the correlation functions.

We start by focusing on the two-point correlations. Note that the anomalous graphs found thus far, in figure 3 and 4, all involve "nearly-closed" loops: the trees have paths that go from the neighborhood of $\vec{i}$ to the neighborhood of $\vec{j}$, and from the neighborhood of $\vec{j}$ to the neighborhood of $\vec{i}$. The paths do not actually form closed loops, since no closed loops are allowed in spanning trees, but they do come very close (within one site). The reasons for this are general, so similar structures will arise in all anomalous graphs, for all correlation functions. For example, the anomalous graphs in figure 4 arose because a long, nearly-closed loop from one site could not be reversed in direction, if it passed through fixed arrows at the other site (see figure 10 of appendix B).

In the open case, these anomalous graphs between $\vec{i}$ and $\vec{j}$ always fall off faster than $\mathcal{O}(1/x^4)$. This is in contrast to the closed boundary case, where such diagrams diverge—see Eqs. (6-8). The difference results from the Green functions. While the Green function diverges as $\ln(x)$ along closed boundaries, it decays as $1/x^2$ along open boundaries (see appendix C). Using Priezzhev's method, the matrix determinant for evaluating any closed loop diagrams necessarily involves two Green functions, one from $\vec{i}$ to $\vec{j}$, and another from $\vec{j}$ to $\vec{i}$, giving an overall factor of $1/x^4$. Furthermore, calculating the diagrams requires two matrix determinants, which come with leading terms equal in magnitude, but opposite in sign—see figure 6. The $\mathcal{O}(1/x^4)$ parts of the closed loop diagrams thus cancel along open boundaries. So the anomalous graphs for the two-point functions automatically fall off faster than $\mathcal{O}(1/x^4)$, and do not need to be considered when calculating leading-order, universal parts of correlation functions.

By this logic, for any n-point open boundary correlations, any anomalous graphs must have terms that decay as $1/(x_a - x_b)^p$, where $p \geq 5$, for some $a, b \in 1, 2, \cdots n$.

Aside from the sites at x_a and x_b, there are $(n-2)$ other sites that need to appear in the connected correlation function. Each brings a new Green function, of $\mathcal{O}(1/x^2)$, so the overall contribution of any anomalous graph must decay at least as fast as $\mathcal{O}(1/x^{5+2(n-2)}) = \mathcal{O}(1/x^{2n+1})$. But we will see in the next section that all n-point correlations decay to leading order as $1/x^{2n}$. So the anomalous graphs have no effect on the universal parts of any n-point correlation functions. The conjecture at the end of section V has thus been proven for all open boundary correlations.

X. ALL n-POINT CORRELATIONS ALONG OPEN BOUNDARIES, PART II

Since we can ignore the anomalous graphs for open boundary correlation functions, no error is introduced by writing the height probability at each site as a linear combination of local arrow diagrams, independently using at each site the linear relationships derived for the one-point functions. Defining the open boundary correlation f_{op} analogously to f_c for the closed case (Eq. (12)), we then have

$$f_{\mathrm{op}}(a_1, a_2, \ldots, a_n) \ =$$

$$= \sum_{u_1=1}^{N_{\mathrm{loc}}} \sum_{u_2=1}^{N_{\mathrm{loc}}} \ldots \sum_{u_n=1}^{N_{\mathrm{loc}}} D_{a_1 u_1} D_{a_2 u_2} \ldots D_{a_n u_n}$$
$$\langle L_{\mathrm{op},u_1}(x_1) L_{\mathrm{op},u_2}(x_2) \ldots L_{\mathrm{op},u_n}(x_n) \rangle \ (30)$$

Each $L_{\mathrm{op},u_f}(x_f)$ represents a local arrow diagram at x_f, analogous to the diagrams in figure 2, but for the open case, and N_{loc} is the total number of possible local arrow diagrams at a single site (see [16] for the list of diagrams). $\mathbf{D}$ is a constant matrix expressing height probabilities in terms of local arrow diagrams, for one-point functions, and was (implicitly) found in [16]. Each correlation of local arrow diagrams can now be calculated with the Majumdar-Dhar method.

If a site $\vec{i}$ has local arrow constraints u, we express those constraints by a matrix $\mathbf{B}_u$, and let $\mathbf{G}_{uu}$ be the Green function matrix for the sites around $\vec{i}$. $\mathbf{B}_u$ and $\mathbf{G}_{uu}$ are both associated only with sites in the vicinity of $\vec{i}$. $p_u = \det(\mathbb{I} + \mathbf{B}_u \mathbf{G}_{uu})$ gives the one-point probability for the local arrow diagram $L_{\mathrm{op},u}$. The two-point correlation of local arrow diagrams u_1 and u_2 is given by $\det(\mathbb{I}+\mathbf{BG})$, where $\mathbf{B}$ is block diagonal, with $\mathbf{B}_{u_1}$ and $\mathbf{B}_{u_2}$ along the block diagonal, and $\mathbf{G}$ is made of the four matrix blocks $\mathbf{G}_{u_1 u_1}$, $\mathbf{G}_{u_1 u_2}$, $\mathbf{G}_{u_2 u_1}$, and $\mathbf{G}_{u_2 u_2}$. Mahieu and Ruelle found that the leading order contribution to the *bulk* two-point probability is given by [15]

$$\det(\mathbb{I} + \mathbf{BG}) = -p_{u_1} p_{u_2} \mathrm{Trace} \left\{ \frac{\mathbb{I}}{\mathbb{I} + \mathbf{B}_{u_1} \mathbf{G}_{u_1 u_1}} \mathbf{B}_{u_1} \mathbf{G}_{u_1 u_2} \frac{\mathbb{I}}{\mathbb{I} + \mathbf{B}_{u_2} \mathbf{G}_{u_2 u_2}} \mathbf{B}_{u_2} \mathbf{G}_{u_2 u_1} \right\} \quad (31)$$

Similarly, they found that the bulk, leading-order, contribution to the three-point probability is given by

$$\det(\mathbb{I} + \mathbf{BG}) \ =$$
$$p_{u_1} p_{u_2} p_{u_3} \mathrm{Trace} \left\{ \frac{\mathbb{I}}{\mathbb{I} + \mathbf{B}_{u_1} \mathbf{G}_{u_1 u_1}} \mathbf{B}_{u_1} \mathbf{G}_{u_1 u_2} \frac{\mathbb{I}}{\mathbb{I} + \mathbf{B}_{u_2} \mathbf{G}_{u_2 u_2}} \mathbf{B}_{u_2} \mathbf{G}_{u_2 u_3} \frac{\mathbb{I}}{\mathbb{I} + \mathbf{B}_{u_3} \mathbf{G}_{u_3 u_3}} \mathbf{B}_{u_3} \mathbf{G}_{u_3 u_1} \right\} +$$
$$p_{u_1} p_{u_2} p_{u_3} \mathrm{Trace} \left\{ \frac{\mathbb{I}}{\mathbb{I} + \mathbf{B}_{u_1} \mathbf{G}_{u_1 u_1}} \mathbf{B}_{u_1} \mathbf{G}_{u_1 u_3} \frac{\mathbb{I}}{\mathbb{I} + \mathbf{B}_{u_3} \mathbf{G}_{u_3 u_3}} \mathbf{B}_{u_3} \mathbf{G}_{u_3 u_2} \frac{\mathbb{I}}{\mathbb{I} + \mathbf{B}_{u_2} \mathbf{G}_{u_2 u_2}} \mathbf{B}_{u_2} \mathbf{G}_{u_2 u_1} \right\}$$
$$(32)$$

(Eqs. (31-32) are written in a different form than the expressions in [15], but are equivalent.)

The derivation in [15] of Eq. (31) in the bulk relied on the fact that the leading-order contribution to the two-point function comes from the pieces of $\det(\mathbb{I}+\mathbf{BG})$ with two terms off the block diagonal (i.e. one term from $\mathbf{G}_{u_1 u_2}$, and one term from $\mathbf{G}_{u_2 u_1}$). Similarly, the derivation of Eq. (32) was based on the fact that the leading-order, connected, contribution to the three-point function comes from the terms of $\det(\mathbb{I} + \mathbf{BG})$ with three terms off the block diagonal.

The trace formulae can be extended for all higher-order correlations for the open case. We will see that the leading-order contribution to the open boundary n-point function decays as $\mathcal{O}(1/x^{2n})$. The open boundary Green function (appendix C) between (x_1, y_1) and (x_2, y_2) is

$$G_{\mathrm{op},0}(x_1, y_1; x_2, y_2) = \frac{(y_1 + 1)(y_2 + 1)}{\pi(x_1 - x_2)^2} + \ldots \quad (33)$$

Here, x labels distance along the boundary, and y labels distance from the boundary (the boundary is at $y = 0$). Since the Green function decays as $1/x^2$, we can only have n terms off the block diagonal. Furthermore, to get

a connected correlation function, we must have exactly one term off the block diagonal in every block row and in every block column. This allows us to generalize Eqs. (31-32) for open boundary n-point functions; they generalize in the obvious manner, with $(n-1)!$ trace terms for the n-point function, corresponding to the $(n-1)!$ ways that we can loop through the n positions.

Eq. (33) shows that each off-diagonal block, $\mathbf{G}_{uv}$, factorizes into the product of a column vector and row vector:

$$\mathbf{G}_{u_f u_g} = \frac{1}{\pi(x_f - x_g)^2}\mathbf{h}_{u_f}\mathbf{h}_{u_g}^T \ , \qquad (34)$$

where $\mathbf{h}_{u_f}$ is a column vector of heights $y+1$ of the sites around x_f in $L_{\text{op},u_f}(x_f)$—i.e. the pth entry of $\mathbf{h}_{u_f}$ is the value of $y+1$ for the pth site of $L_{\text{op},u_f}(x_f)$. Substituting this in the generalization of Eqs. (31-32), and using the cyclicity of the trace, each of the $(n-1)!$ matrix traces becomes a product of n 1×1 matrices. Furthermore, the $(n-1)!$ traces differ from each other only in the $1/(x_f - x_g)^2$ terms chosen. The leading-order, connected part of the correlation function of n local arrow diagrams is then found to be

$$\langle L_{\text{op},u_1}(x_1)L_{\text{op},u_2}(x_2)\ldots L_{\text{op},u_n}(x_n)\rangle =$$
$$= \left(\prod_{f=1}^{n} k_{u_f}\right) \det \mathbf{M} \quad (35)$$

$\mathbf{M}$ is defined as the $n \times n$ matrix

$$M_{fg} \equiv \begin{cases} 0 & \text{if } f = g \\ 1/(x_f - x_g)^2 & \text{if } f \neq g \end{cases} \qquad (36)$$

and the k_u are simply numbers:

$$k_u \equiv \frac{1}{\pi} \det\left(\mathbb{I} + \mathbf{B}_u\mathbf{G}_{uu}\right)\left(\mathbf{h}_u^T \frac{\mathbb{I}}{\mathbb{I} + \mathbf{B}_u\mathbf{G}_{uu}}\mathbf{B}_u\mathbf{h}_u\right) \quad (37)$$

Inserting this into Eq. (30) gives all open boundary n-point correlations. To express our results in a simpler manner, we define

$$\phi_a(x) \equiv \frac{\delta_{h_x,a} - p_{a,\text{op}}}{K_a} \qquad , \text{where a} = 1,\ldots 4 \qquad (38)$$

We have defined the following constants:

$$
\begin{aligned}
&p_{1,\text{op}} = \frac{9}{2} - \frac{42}{\pi} + \frac{320}{3\pi^2} - \frac{512}{9\pi^3} && K_1 = -\frac{3}{\pi} + \frac{80}{3\pi^2} - \frac{512}{9\pi^3} \\
&p_{2,\text{op}} = -\frac{33}{4} + \frac{66}{\pi} - \frac{160}{\pi^2} + \frac{1024}{9\pi^3} && K_2 = \frac{9}{\pi} - \frac{200}{3\pi^2} + \frac{1024}{9\pi^3} \\
&p_{3,\text{op}} = \frac{15}{4} - \frac{22}{\pi} + \frac{160}{3\pi^2} - \frac{512}{9\pi^3} && K_3 = -\frac{7}{\pi} + \frac{40}{\pi^2} - \frac{512}{9\pi^3} \\
&p_{4,\text{op}} = 1 - \frac{2}{\pi} && K_4 = \frac{1}{\pi}
\end{aligned}
\qquad (39)
$$

$p_{a,\text{op}}$ is the probability for a site along an open boundary to have height a, and the K_a are normalization factors. We then, finally, have

$$\langle \phi_{a_1}(x_1)\phi_{a_2}(x_2)\ldots\phi_{a_n}(x_n)\rangle = \det(\mathbf{M}) \qquad (40)$$

For $n = 2$, this reproduces the open boundary one- and two-point functions found in [16].

$\det(\mathbf{M})$ is the same as the n-point function of $-2\partial\theta\partial\bar{\theta}$, so up to rescaling factors ($-2K_a$'s), all four height variables are represented by $\partial\theta\partial\bar{\theta}$ along open boundaries. This is rather surprising, given that we have seen that the height variables are represented by different operators along closed boundaries (Eqs. (20-22)). In CFT's, boundary operators are derived from OPE's of bulk operators—so the fact that the height operators are different along closed boundaries proves that they must be different in the bulk, but apparently these different bulk operators become identical along open boundaries.

We nowhere used the fact that these were the local arrow diagrams associated with the height variables. So, in fact, we have shown that *all* local arrow diagrams along open boundaries are represented by $\partial\theta\partial\bar{\theta}$.

We have also found the correlation function of n unit height variables along *closed* boundaries. This requires local arrow constraints at $3n$ vertices of the ASM, and thus the calculation of a $3n$-dimensional matrix determinant. The matrix is divided into 3 by 3 block submatrices, such that the diagonal blocks are all identical, and the off-diagonal blocks all have the same form. A rotation makes the matrix diagonal in 2 out of every 3 rows (and columns). The universal part of the correlation function is thus found to be

$$\left(\frac{3\pi - 8}{\pi^2}\right)^n \det\mathbf{M} \qquad (41)$$

This confirms the field identification in Eq. (20).

XI. n-POINT CORRELATIONS ALONG OPEN BOUNDARIES, WITH DISSIPATIVE DEFECTS

Along open boundaries, the defect-free Green function, $\mathbf{G}_0 = \mathbf{G}_{\text{op},0}$ does not diverge as $L \to \infty$, so for a single dissipative defect we can modify the Green function as in Eq. (25). Using this new Green function, the open height probabilities at $(x_1, 0)$, for a defect of strength k at $\vec{d} = (0, y)$ are

$$\text{f}_{\text{op}}(a) = -K_a \frac{k(y+1)^2}{\pi(1 + kG_{\text{op},0}(\vec{d},\vec{d}))}\frac{1}{x_1^4}, \quad a = 1,2,3,4 \qquad (42)$$

The same K_a factors that we saw in the height-height correlations appear in height-defect correlations.

We define an operator $\phi_{5;k}(\vec{d})$, corresponding to the addition of a defect of strength k at $\vec{d} = (x,y)$, and then multiplication of all correlation functions by a normalization factor

$$\frac{\pi(1 + kG_{\text{op},0}(\vec{d}, \vec{d}))}{k(y+1)^2} \qquad (43)$$

Then Eq. (42) becomes

$$\langle \phi_a(x_1)\phi_{5;k}(x_2) \rangle = -\frac{1}{(x_1 - x_2)^4}, \quad a \neq 5 \qquad (44)$$

$\phi_{5;k}$ acts just like any of the four height variables in two-point correlations (Eq. (44) is Eq. (40) with $n = 2$). In fact, we find that $\phi_{5;k}$ acts like ϕ_1, ϕ_2, ϕ_3, and ϕ_4 in all higher-order correlation functions, containing multiple height variables and multiple dissipative defects.

Suppose we are calculating a correlation function with n height variables, and m dissipative defects. The dissipative defects are at $\vec{d}_w = (x_w, y_w)$, and have strength k_w, $1 \leq w \leq m$. As with the height locations, the x_w coordinates of the defects all scale with the same factor x, where $x \to \infty$. The change in the toppling matrix , $\delta\mathbf{\Delta} \equiv \mathbf{\Delta} - \mathbf{\Delta}_0$, is

$$\delta\Delta_{\vec{i},\vec{j}} = \begin{cases} k_w & \text{if } \vec{i} = \vec{j} = \vec{d}_w, \quad 1 \leq w \leq m \\ 0 & \text{otherwise} \end{cases} \qquad (45)$$

The Green function is modified from its defect-free value, $\mathbf{G}_{\text{op},0}$, to

$$\begin{aligned} \mathbf{G} &= \frac{\mathbb{I}}{\mathbf{\Delta}} = \frac{\mathbb{I}}{\mathbf{\Delta}_0 + \delta\mathbf{\Delta}} = \frac{\mathbf{G}_{\text{op},0}}{\mathbb{I} + (\delta\mathbf{\Delta})\mathbf{G}_{\text{op},0}} \\ &= \sum_{p=0}^{\infty} \mathbf{G}_{\text{op},0}\left(-(\delta\mathbf{\Delta})\mathbf{G}_{\text{op},0}\right)^p \end{aligned} \qquad (46)$$

$G(\vec{i}, \vec{j})$ can be represented as a trip from $\vec{i}$ to $\vec{j}$, where along the trip, the traveller can visit any of the defect sites as often as he or she wishes, each time picking up a factor of $-(\delta\mathbf{\Delta})\mathbf{G}_{\text{op},0}$.

We have already seen that the defect-free correlation function of n height variables has a leading term of $\mathcal{O}(1/x^{2n})$. If we instead use the Green function with defects, each trip to a defect introduces a factor of $1/x^2$ (see Eq. (33)). In a connected function, we should visit each defect at least once; in the leading term, each defect will be visited from a distant site exactly once, and the correlation function will have a leading term of $\mathcal{O}(1/x^{2(n+m)})$.

After visiting $\vec{d}_w$, we may travel repeatedly from $\vec{d}_w$ to $\vec{d}_w$ without picking up extra factors of $1/x^2$. This produces a contribution to Eq. (46) of

$$\sum_{p=0}^{\infty} (-k_w G_0(\vec{d}_w, \vec{d}_w))^p = \frac{1}{1 + k_w G_0(\vec{d}_w, \vec{d}_w)} \qquad (47)$$

(We already saw this factor for a single dissipative defect, in Eq. (25).) Furthermore, inspection of Eq. (33) shows that the visit to the defect at $\vec{d}_w$ from another site will result in a factor of $k_w(y_w + 1)^2/\pi$. With Eq. (47), this motivates the normalization factor in Eq. (43).

Eq. (43) normalizes the correlation function of n height variables and m defects. To see that the form of the correlation function is still $\det(\mathbf{M})$, expand the determinant out into cycles. The connected part of the determinant in Eq. (40) is a sum of closed cycles of length n, where each cycle visits each of the positions $(x_f, 0)$ exactly once, and picks up a factor of $1/(x_f - x_g)^2$ when it travels from x_f to x_g. After normalizing, the defects have exactly the same effect as the height variables—each trip to (or from) a defect results in a $1/x^2$ term from the Green function to (or from) the defect (Eq. (33)).

So, in the end, the correlation function of n height variables on the boundary and m defect sites near or on the boundary, is given by the $(m + n)$ dimensional matrix determinant, $\det(\mathbf{M})$ (with appropriate normalization factors). This shows that dissipative defect sites along or near open boundaries are, like the height variables, represented by $\partial\theta\partial\bar{\theta}$.

Note that a dissipative defect has a much larger effect along a closed boundary than along an open one. A defect is represented by a dimension zero operator along a closed boundary, but by a dimension two operator along an open boundary. This makes sense; along open boundaries, grains are already dissipated by topplings, so adding a little more dissipation has only minor effects, compared to dissipation on a closed boundary.

APPENDIX A: ANOMALOUS GRAPHS IN BOUNDARY TWO-POINT CORRELATIONS—PART I

In this section we discuss what we call anomalous graphs of the first kind, which arise when converting from ASM height probabilities to spanning tree probabilities. As stated in section V, it would be natural to expect, based on analogy with the one-point height probabilities, for the two-point height probabilities to be given by Eq. (5). However, this turns out to not be the case. Let us carefully consider how height correlations can be
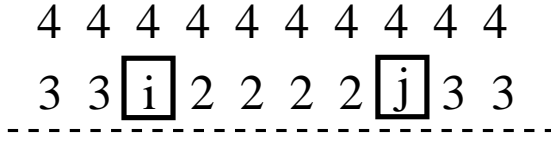
$$4\ 4\ 4\ 4\ 4\ 4\ 4\ 4\ 4\ 4$$

$$3\ 3\ \boxed{i}\ 2\ 2\ 2\ 2\ \boxed{j}\ 3\ 3$$

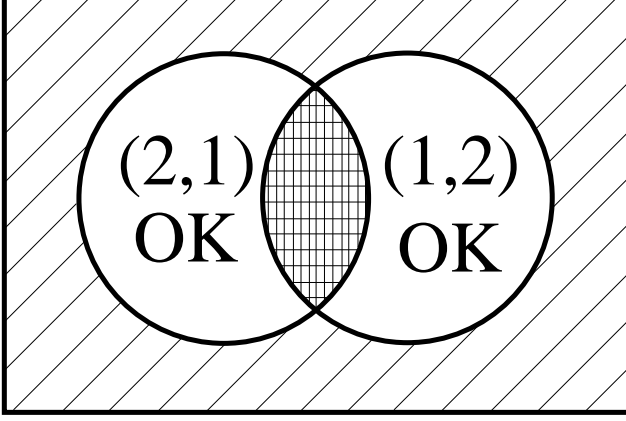FIG. 7: State not in any S_{kl}, and in multiple $\tilde{S}_{kl}$.

FIG. 8: Venn diagram in the space of states where $(h_{\vec{i}}, h_{\vec{j}}) = (2,2)$ is allowed, but $(h_{\vec{i}}, h_{\vec{j}}) = (1,1)$ is not.

turned into spanning tree probabilities. We focus on the closed two-point correlations; other cases are similar.

For correlations between $\vec{i}$ and $\vec{j}$, Ivashkevich divided the states of the ASM into sets S_{kl}, consisting of states allowed when $h_{\vec{i}} \geq k$ and $h_{\vec{j}} \geq l$, but forbidden otherwise [16]. However, not all ASM states fall into one of these sets. There are states that are allowed when $(h_{\vec{i}}, h_{\vec{j}}) = (1,2)$, and when $(h_{\vec{i}}, h_{\vec{j}}) = (2,1)$, but forbidden when $(h_{\vec{i}}, h_{\vec{j}}) = (1,1)$; these states do not belong to any set S_{kl}. One such state is shown in figure 7.

We find it convenient to define $\tilde{S}_{kl}$, consisting of ASM height configurations on the sandpile, *excluding* $\vec{i}$ and $\vec{j}$, that are allowed when we add $(h_{\vec{i}}, h_{\vec{j}}) = (k, l)$, but both forbidden when we add $(h_{\vec{i}}, h_{\vec{j}}) = (k-1, l)$, and also forbidden when we add $(h_{\vec{i}}, h_{\vec{j}}) = (k, l-1)$. Note that when counting states, the fact that we do not specify the heights of $\vec{i}$ and $\vec{j}$ introduces a multiplicative factor; for example, $|\tilde{S}_{12}| = |S_{12}|/(3*2)$ for $\vec{i}$ and $\vec{j}$ both on a closed boundary. Now, every state must be in at least one of the sets $\tilde{S}_{kl}$, but some states are in several $\tilde{S}_{kl}$'s. For example, the state in figure 7 is in both $\tilde{S}_{12}$ and $\tilde{S}_{21}$.

No anomalous graphs arise for two-point correlations involving at least one unit height variable, since in those cases the relevant $\tilde{S}$-sets do not intersect. The number of states where $\vec{i}$ has height one and $\vec{j}$ has height h is $\sum_{p=1}^{h} |\tilde{S}_{1p}|$. Furthermore, the representation of $\tilde{S}_{1h}$ is exactly what we would expect; it corresponds to the number of spanning trees where $\text{NNP}_{\vec{i}} = 0$, and $\text{NNP}_{\vec{j}} = h - 1$. So no anomalous graphs arise when the two-point correlation has at least one unit height variable.

Things get more complicated when both heights are higher heights. We discuss in detail the 2-2 (height two-height two) correlation along a closed boundary; the analysis for the other two-point correlations is similar.

If $\vec{i}$ and $\vec{j}$ both have height two, we must be in at least one $\tilde{S}_{kl}$, for $k \leq 2$, $l \leq 2$. $\tilde{S}_{12}$ and $\tilde{S}_{21}$ intersect, so number of 2-2 states is

$$|\tilde{S}_{11}| + |\tilde{S}_{12}| + |\tilde{S}_{21}| + \left(|\tilde{S}_{22}| - |\tilde{S}_{12} \cap \tilde{S}_{21}|\right) \tag{A1}$$

The first three terms all have the "natural" spanning tree representation. The difficulty is in evaluating $|\tilde{S}_{22}| - |\tilde{S}_{12} \cap \tilde{S}_{21}|$. $\tilde{S}_{22}$ counts states where $(h_{\vec{i}}, h_{\vec{j}}) = (2,2)$ is allowed, but neither $(h_{\vec{i}}, h_{\vec{j}}) = (1,2)$ nor $(h_{\vec{i}}, h_{\vec{j}}) = (2,1)$ are allowed. In figure 8, we have started with a large rectangle, representing the the set of states where $(h_{\vec{i}}, h_{\vec{j}}) = (2,2)$ is allowed, and $(h_{\vec{i}}, h_{\vec{j}}) = (1,1)$ is forbidden—we call this set X. In the rectangle are two subsets, corresponding to regions where $(h_{\vec{i}}, h_{\vec{j}}) = (1,2)$ is allowed, and where $(h_{\vec{i}}, h_{\vec{j}}) = (2,1)$ is allowed—we call these two sets X_1 and X_2. In this Venn diagram, $\tilde{S}_{22}$ is the diagonally shaded region outside the circles, and $\tilde{S}_{12} \cap \tilde{S}_{21}$ is the cross-hatched intersection of the two circles. Looking at the Venn diagram, we see that to find $|\tilde{S}_{22}| - |\tilde{S}_{12} \cap \tilde{S}_{21}|$, we start with the set X, and then subtract off the states in X_1 and X_2 *independently*. By independently we mean that states in the intersection of X_1 and X_2 get subtracted off twice. So

$$|\tilde{S}_{22}| - |\tilde{S}_{12} \cap \tilde{S}_{21}| = |X| - |X_1| - |X_2| \tag{A2}$$

For states in X, setting $(h_{\vec{i}}, h_{\vec{j}}) = (1,1)$ produces an MFSC. (Note that we have defined the MFSC as the maximal FSC produced when the heights at both $\vec{i}$ and $\vec{j}$ are simultaneously set to one; if only one height was set to one, then the largest FSC might be smaller, or there might be no FSC at all.) The set X can be partitioned into the following disjoint subsets, depending on the shape of the MFSC:

X_A : The MFSC consists of disjoint subsets around $\vec{i}$ and $\vec{j}$

X_B : The MFSC is connected, with ≥ 2 neighbors of $\vec{i}$ and exactly 1 neighbor of $\vec{j}$

X_C : The MFSC is connected, with ≥ 2 neighbors of $\vec{j}$ and exactly 1 neighbor of $\vec{i}$

X_D : The MFSC is connected, with exactly 1 neighbor of $\vec{\mathrm{i}}$ and exactly 1 neighbor of $\vec{\mathrm{j}}$

X_1 and X_2 can be partitioned into analogously defined subsets, X_{1A}, X_{1B}, etc. . . So, for example, X_{1D} is the subset of X_D such that $(h_{\vec{i}}, h_{\vec{j}}) = (1, 2)$ is allowed. (Note that $X_{1C} = \emptyset$ and $X_{2B} = \emptyset$.) We want

$$\sum_{k \in \{A,B,C,D\}} (\mid X_k \mid - \mid X_{1k} \mid - \mid X_{2k} \mid) \qquad \text{(A3)}$$

Our "naive" guess would be that this would equal the set of states where $\mathrm{NNP}_{\vec{i}} = \mathrm{NNP}_{\vec{j}} = 1$. We carefully count the states, comparing with this guess.

X_A, after subtracting off the states from X_{1A} and X_{2A}, is equal to one-fourth the number of spanning trees for which $\vec{i}$ and $\vec{j}$ each have one NNP, and $\{\vec{i}\}\cup\mathrm{Tree}_{\vec{i}}$ does not border or intersect $\{\vec{j}\} \cup \mathrm{Tree}_{\vec{j}}$. (The one-fourth comes from the fact that the spanning tree arrows from $\vec{i}$ and $\vec{j}$ can point out from the MFSC in any direction.) $\mathrm{Tree}_{\vec{i}}$ refers to the set of sites that are predecessors of i. The condition that $\{\vec{i}\} \cup \mathrm{Tree}_{\vec{i}}$ and $\{\vec{j}\} \cup \mathrm{Tree}_{\vec{j}}$ cannot border each other comes from the condition that the MFSC consist of disjoint subsets around $\vec{i}$ and $\vec{j}$.

We now consider X_B. The MFSC generated when $(h_{\vec{i}}, h_{\vec{j}}) = (1, 1)$ must still be an MFSC when $(h_{\vec{i}}, h_{\vec{j}}) = (2, 1)$. So $(h_{\vec{i}}, h_{\vec{j}}) = (2, 2) \to (2, 1)$ produces an MFSC that includes exactly one neighbor of $\vec{j}$, the site $\vec{i}$, and at least two of $\vec{i}$'s neighbors. Just as in section III, this is equivalent to a modified ASM, where bonds along the border of the MFSC are removed (except for one bond of $\vec{j}$). In Eq. (A3), $X_{2B} = \emptyset$, but we do need to subtract off states in X_{1B}. To do this we only count the states of X_B such that $(h_{\vec{i}}, h_{\vec{j}}) = (1, 2)$ is forbidden, which implies that $h_{\vec{i}} = 2 \to 1$ should produce a new, smaller FSC, completely contained within the larger MFSC. We then see that $\mid X_B \mid - \mid X_{1B} \mid - \mid X_{2B} \mid$ corresponds to one-fourth of all spanning trees where $\mathrm{NNP}_{\vec{i}} = \mathrm{NNP}_{\vec{j}} = 1$, and either $\vec{i} \in \mathrm{Tree}_{\vec{j}}$ or $\vec{i}$ borders $\mathrm{Tree}_{\vec{j}}$, *with one exception.* The exception occurs because the MFSC, by construction of X_B, can only have one neighbor of $\vec{j}$. So $\mid X_B \mid - \mid X_{1B} \mid - \mid X_{2B} \mid$ will not count cases where $\mathrm{NNP}_{\vec{i}} = \mathrm{NNP}_{\vec{j}} = 1$, $\vec{i}$ borders $\mathrm{Tree}_{\vec{j}}$, and $\vec{j}$ borders $\mathrm{Tree}_{\vec{i}}$—this case is shown in figure 3. We label this set of graphs as c. Since c has $\mathrm{NNP}_{\vec{i}} = \mathrm{NNP}_{\vec{j}} = 1$, it would be natural to expect it to appear in the spanning trees contributing to the height two-height two correlation function. However, since no MFSC's of X have two neighbors of $\vec{i}$ and two neighbors of $\vec{j}$, graph c does not appear in X_C or X_D either.

The analysis for X_C is identical to that for X_B, and $\mid X_C \mid - \mid X_{1C} \mid - \mid X_{2C} \mid$ counts one-fourth the spanning trees where $\mathrm{NNP}_{\vec{i}} = \mathrm{NNP}_{\vec{j}} = 1$, and either $\vec{j} \in \mathrm{Tree}_{\vec{i}}$ or $\vec{j}$ borders $\mathrm{Tree}_{\vec{i}}$, except that, again, the spanning trees of c are excluded.

FIG. 9: Correlations where $\beta_1 \neq \beta_2$

In X_D, the MFSC has only one neighbor of $\vec{i}$ and one neighbor of $\vec{j}$. The one-to-one mapping between X_D and spanning tree states posessess some subtleties, but the end result is what one would expect: The number of states in $\mid X_D \mid - \mid X_{1D} \mid - \mid X_{2D} \mid$ is one-fourth the number of spanning trees where $\mathrm{NNP}_{\vec{i}} = \mathrm{NNP}_{\vec{j}} = 1$, and $\mathrm{Tree}_{\vec{i}}$ and $\mathrm{Tree}_{\vec{j}}$ border each other, but $\vec{i} \notin \mathrm{Tree}_{\vec{j}}$, and $\vec{j} \notin \mathrm{Tree}_{\vec{i}}$.

In the end, we see that $4 \left(\mid \tilde{S}_{22} \mid - \mid \tilde{S}_{12} \cap \tilde{S}_{21} \mid \right)$ is equal to the number of spanning trees where $\mathrm{NNP}_{\vec{i}} = \mathrm{NNP}_{\vec{j}} = 1$, except for the set c, which contains all spanning trees where both $\vec{i}$ borders $\mathrm{Tree}_{\vec{j}}$, and $\vec{j}$ borders $\mathrm{Tree}_{\vec{i}}$. c consists of the anomalous graphs of the first kind.

An similar analysis for other closed correlation functions shows that the spanning trees in c contribute to the height two-height three correlation, and get subtracted from the height three-height three correlation (relative to a "naive" approach). These changes are necessary for the height probabilities to all sum to one, so this provides a check on our mapping between ASM states and spanning tree states. In the next appendix, we consider yet another complication that arises in the calculation of the two-point functions.

APPENDIX B: ANOMALOUS GRAPHS IN BOUNDARY TWO-POINT CORRELATIONS—PART II

We saw in the previous appendix that Eq. (5) does not quite hold, but is only off by the anomalous graph c. So except for this complication, the two-point height correlations can be turned into linear combinations of probabilities for spanning trees with nonlocal conditions on $\mathrm{NNP}_{\vec{i}}$ and $\mathrm{NNP}_{\vec{j}}$. As in the previous appendix, we discuss only the closed boundary two-point functions; other cases are similar.

As discussed in section IV, we can write the probability $\mathcal{N}_h$ to have $\mathrm{NNP}_{\vec{i}} = h - 1$ as a linear combination of nonlocal arrow diagrams, which we can then rewrite as

FIG. 10: Correlations where $\phi_1 \neq \phi_2$

follows (see figures 1 and 2):

$$
\begin{aligned}
\mathcal{N}_2 &= 2(\beta_1 + \beta_2 + \gamma) \\
&= 2(-L_{c,1} + 3L_{c,2} + L_{c,3}) + \\
&\qquad 4(\beta_1 - \beta_2) + 2(\phi_1 - \phi_2) \qquad \text{(B1)} \\
\mathcal{N}_3 &= 2(\phi_1 + \phi_2 + \epsilon) + \delta \\
&= 2(2L_{c,1} - 2L_{c,2} + L_{c,5}) + L_{c,4} - \\
&\qquad 4(\beta_1 - \beta_2) - 2(\phi_1 - \phi_2) \qquad \text{(B2)}
\end{aligned}
$$

These relationships hold regardless of the correlation functions that $\vec{i}$ are in. We define operators corresponding to the *local* parts of these terms:

$$
\begin{aligned}
L_{N2} &\equiv 2(-L_{c,1} + 3L_{c,2} + L_{c,3}) & \text{(B3)} \\
L_{N3} &\equiv 2(2L_{c,1} - 2L_{c,2} + L_{c,5}) + L_{c,4} & \text{(B4)}
\end{aligned}
$$

In one-point functions, $\beta_1 = \beta_2$, and $\phi_1 = \phi_2$, so $\mathcal{N}_2$ and $\mathcal{N}_3$ simply become L_{N2} and L_{N3}, which are local, and whose expectation values can be found with the Majumdar-Dhar method, as discussed in section IV. (These are the same relations found in [16].)

Unfortunately, things become more complicated with the ϕ_1 and ϕ_2 terms. If the long path of ϕ_1 avoids arrow restrictions at other sites, the long path can be reversed, and ϕ_1 will be equivalent to ϕ_2. However, if the long path goes through arrow restrictions at other sites of the correlation function, then ϕ_1 will not be equivalent to ϕ_2. Figure 10 compares diagrams that arise in (ϕ_1, ϕ_1) correlations, with diagrams that arise in (ϕ_2, ϕ_2) correlations. Three of the four diagrams shown (labeled with "a"'s) are equivalent, but the fourth one (labeled with "b") is not. The resultant anomalous graphs of the second kind were shown and discussed earlier, in section V (see figure 4). When all the correlations of the $(\phi_1 - \phi_2)$'s in Eqs. (B1-B2) are considered, we find the following for

It would be simplest if in correlations of $\mathcal{N}_2$ and $\mathcal{N}_3$, we could replace $\mathcal{N}_2$ and $\mathcal{N}_3$ with L_{N2} and L_{N3}, since local correlation functions are easily calclated. As in the previous appendix, we call this approach "naive,"—this naive approach does not quite work, and we call the deviations of the correct answers from the naive approach "anomalous graphs of the second kind."

We no longer have $\beta_1 = \beta_2$ and $\phi_1 = \phi_2$ in correlations of $\mathcal{N}_2$ and $\mathcal{N}_3$, because in correlations between distant sites $\vec{i}$ and $\vec{j}$, switching arrows at $\vec{i}$ can affect predecessor relationships at $\vec{j}$. To analyze the $(\beta_1 - \beta_2)$ terms in Eqs. (B1-B2), consider the configuration in figure 9, where we have not specified the direction of the arrow from $\vec{i}$. If the arrow from $\vec{i}$ points to $\vec{i}_1$, then $\vec{j}_1$ is not a predecessor of $\vec{j}$, so the configuration at $\vec{j}$ is γ. And if the arrow from $\vec{i}$ points to $\vec{i}_2$, then $\vec{j}_1$ is a predecessor of $\vec{j}$, so the configuration at $\vec{j}$ is ϕ_2. So switching from β_1 to β_2 at $\vec{i}$ can affect whether the configuration at $\vec{j}$ is ϕ_2 or γ. However, this inequivalence between β_1 and β_2 turns out to have no effect on any correlation functions, to any order, since ϕ_2 and γ always appear in the combination $\phi_2 + \gamma$, in L_{N2} and L_{N3}, and $\phi_2 + \gamma$ has no correlations with $\beta_1 - \beta_2$.

closed boundary correlations:

$$
\begin{aligned}
\langle \mathcal{N}_2(r)\mathcal{N}_2(0) \rangle &= \langle L_{N2}(r)L_{N2}(0) \rangle + 4(a - b) & \text{(B5)} \\
\langle \mathcal{N}_2(r)\mathcal{N}_3(0) \rangle &= \langle L_{N2}(r)L_{N3}(0) \rangle - 4(a - b) & \text{(B6)} \\
\langle \mathcal{N}_3(r)\mathcal{N}_3(0) \rangle &= \langle L_{N3}(r)L_{N3}(0) \rangle + 4(a - b) & \text{(B7)}
\end{aligned}
$$

The correlations involving the height one variable are unaffected by these complications.

To summarize, naively transforming from nonlocal arrow diagrams to local arrow diagrams *independently* at every site of a two-point correlation results in anomalous graphs a and b.

APPENDIX C: GREEN FUNCTIONS

The inverse of the bulk toppling matrix $\boldsymbol{\Delta}_0$ is the lattice Green function, which has long been known [24]. It is given by

$$G_0(\tilde{x}, \tilde{y}) = \int_0^{2\pi} \frac{dp_1}{2\pi} \int_0^{2\pi} \frac{dp_2}{2\pi} \frac{e^{ip_1\tilde{x}+ip_2\tilde{y}}}{4 - 2\cos p_1 - 2\cos p_2} \tag{C1}$$

This integral is divergent, producing terms of order $\ln L$, where L is the system size, but these divergences are usually unimportant, since we are typically concerned with differences in Green functions. For large $\tilde{x}$, $\tilde{y}$ this has the expansion [24]

$$G_0(\tilde{x}, \tilde{y}) = -\frac{1}{4\pi} \ln(\tilde{x}^2 + \tilde{y}^2) - \frac{\gamma}{2\pi} - \frac{\ln 8}{4\pi} + \dots , \tag{C2}$$

where $\gamma = 0.57721\dots$ is the Euler-Mascheroni constant. For sites (x_1, y_1) and (x_2, y_2) near an open boundary, where x is the coordinate along the boundary, and y is the distance from the boundary (located at $y = 0$), the Green function is [25]

$$G_{\text{op},0}(x_1, y_1; x_2, y_2) = G_0(x_1 - x_2, y_1 - y_2) - \\ G_0(x_1 - x_2, y_1 + y_2 + 2) \tag{C3}$$

Along closed boundaries, it is [25]

$$G_{\text{c},0}(x_1, y_1; x_2, y_2) = G_0(x_1 - x_2, y_1 - y_2) + \\ G_0(x_1 - x_2, y_1 + y_2 + 1) \tag{C4}$$

The minus sign between the Green functions in Eq. (C3) cancels out divergences in the Green function. The expansion of the Green function for points along the boundary has already been calculated [25], and can be extended to points near the boundary, but far from each other ($y_1 = \mathcal{O}(1)$, $y_2 = \mathcal{O}(1)$, $| x_1 - x_2 |\to \infty$), by the recursion relationships, $\mathbf{G}_0 \boldsymbol{\Delta}_0 = \mathbb{I}$. We find

$$G_{\text{op}}(x_1, y_1; x_2, y_2) = \frac{(y_1 + 1)(y_2 + 1)}{\pi(x_1 - x_2)^2} + \dots \tag{C5}$$

and

$$G_{\text{c}}(x_1, y_1, x_2, y_2) = -\frac{1}{\pi} \ln | x_1 - x_2 | -(\frac{\gamma}{\pi} + \frac{3\ln 2}{2\pi}) - (3y_1(y_1 + 1) + 3y_2(y_2 + 1) + 1)\frac{1}{6\pi(x_1 - x_2)^2}$$
$$+ \left(y_1(y_1 + 1)(y_1^2 + y_1 - 1) + y_2(y_2 + 1)(y_2^2 + y_2 - 1) + 6y_1(y_1 + 1)y_2(y_2 + 1) - \frac{17}{60} \right) \frac{1}{4\pi(x_1 - x_2)^4} + \dots \tag{C6}$$

APPENDIX D: MORE CLOSED BOUNDARY CORRELATION FUNCTIONS

Here we list the three-point correlation functions along closed boundaries that were not stated in section VI. As a check, the correlation functions in this appendix were found by the methods already described in sections V and VI. However, they can all also be determined from those already listed in section VI, from the requirement that the three height probabilities must sum at all sites, and by symmetry. They are listed here only for reference, and because they provide checks on our calculations. We find

$$f_{\text{c}}(1, 1, 3) = -\frac{2(4 - \pi)(3\pi - 8)^2}{\pi^6(x_1 - x_2)^2(x_1 - x_3)^2(x_2 - x_3)^2} \\ + \frac{(3\pi - 8)^2}{\pi^5(x_1 - x_3)^3(x_2 - x_3)^3} + \dots \tag{D1}$$

$$f_{\text{c}}(1, 2, 3) = \frac{4(\pi - 3)(\pi + 8)(3\pi - 8)}{\pi^6(x_1 - x_2)^2(x_1 - x_3)^2(x_2 - x_3)^2}$$

$$-\frac{(3\pi - 8)^2}{\pi^5(x_1 - x_2)^3(x_2 - x_3)^3} + \\ -\frac{(3\pi - 8)(24 - 7\pi)}{2\pi^5(x_1 - x_2)^3(x_1 - x_3)^3} \tag{D2}$$

$$f_{\text{c}}(1, 3, 3) = \frac{(3\pi - 8)(4 - \pi)(8 + \pi)}{\pi^6(x_1 - x_2)^2(x_1 - x_3)^2(x_2 - x_3)^2}$$
$$+ \frac{(3\pi - 8)(8 - \pi)}{2\pi^5(x_1 - x_2)^3(x_1 - x_3)^3} + \dots \tag{D3}$$

We can now check that $f_c(1, 1, 1) + f_c(1, 1, 2) + f_c(1, 1, 3) = 0$, as it must. Interchanging x_2 and x_3 in $f_c(1, 1, 2)$ gives

$$f_{\text{c}}(1, 2, 1) = -\frac{8(\pi - 3)(3\pi - 8)^2}{\pi^6(x_1 - x_2)^2(x_1 - x_3)^2(x_2 - x_3)^2}$$
$$-\frac{(3\pi - 8)^2}{\pi^5(x_1 - x_2)^3(x_3 - x_2)^3} + \dots \tag{D4}$$

We can then check that $f_c(1, 1, 1) + f_c(1, 2, 1) + f_c(1, 3, 1) = 0$. Three-point correlation functions with no unit height variables *cannot* be found with the methods in

this paper, as already discussed in section VI. However, if we use the conjecture proposed in section V, of dropping the anomalous graphs (as we did to obtain Eq. (23)), we now obtain

$$f_c(2,2,3) = \frac{(24-5\pi)(-192+112\pi-13\pi^2)}{4\pi^6(x_1-x_2)^2(x_1-x_3)^2(x_2-x_3)^2}$$
$$+\frac{(3\pi-8)(7\pi-24)}{2\pi^5(x_1-x_3)^3(x_2-x_3)^3}+\cdots \quad \text{(D5)}$$

$$f_c(2,3,3) = -\frac{(8+\pi)(-192+112\pi-13\pi^2)}{4\pi^6(x_1-x_2)^2(x_1-x_3)^2(x_2-x_3)^2}$$
$$-\frac{(3\pi-8)(8-\pi)}{2\pi^5(x_1-x_2)^3(x_1-x_3)^3}+\cdots \quad \text{(D6)}$$

$$f_c(3,3,3) = -\frac{(8+\pi)(64-32\pi+\pi^2)}{4\pi^6(x_1-x_2)^2(x_1-x_3)^2(x_2-x_3)^2}+\cdots, \quad \text{(D7)}$$

As with the other correlation functions in this section, these agree with the requirements that the three height probabilities must sum to one at all sites, and with the field identifications in Eqs. (20-22).

APPENDIX E: THE $c=-2$ CONFORMAL FIELD THEORY

The central charge -2 conformal field theory is perhaps the simplest known logarithmic conformal field theory. While the theory has a simple underlying Gaussian structure, it still possesses a number of subtle features. We use the formulation of the $c=-2$ CFT where the action is given by

$$S = \frac{1}{\pi}\int \partial\theta\,\bar\partial\bar\theta \quad \text{(E1)}$$

∂ and $\bar\partial$ refer to the holomorphic and antiholomorphic derivatives—that is, the derivatives with respect to $z = x+iy$ and $\bar z = x-iy$. θ and $\bar\theta$ are anticommuting Grassmanian variables. The action has zero modes, which make the partition function zero. If we normalize the action by not integrating over the zero modes, we get Wick contraction rules, with each contraction between $\theta(z_1)$ and $\bar\theta(z_2)$ giving a factor of $-\log(z_1-z_2)$.

While boundary conformal field theories are generally well understood [19, 28], boundary logarithmic conformal field theories possess a number of subtleties that are not well understood. Results on boundary LCFT are still to some degree contradictory [29–34]. However, basic results from non-logarithmic boundary CFT

should still be expected to apply. In particular, just as for non-logarithmic boundary CFT's, as bulk operators are moved near a boundary, their antiholomorphic pieces should behave like holomorphic pieces at mirror locations across the boundary [19, 30].

APPENDIX F: A PROOF THAT THE HEIGHT VARIABLES HAVE DIFFERENT BULK FIELD IDENTIFICATIONS

The correlation functions in Eqs. (13-15), Eqs. (17-19), and Eqs. (27-29) show conclusively that the three height variables are represented by different operators along closed boundaries. As already discussed, since boundary operators are derived from OPE's of bulk operators, this proves that the height variables must be represented by different operators in the bulk as well [19]. However, it is worth noting that this conclusion can be reached with a simple argument, based on general principles of conformal field theory, without doing any detailed calculations.

Suppose that all four height variables were represented (up to multiplicative factors) by the same field operator. The unit height variable is known, from its two-point correlation, to have dimension two [14], so, by our assumption, all four height variables would have scaling dimension two. The height probabilities get modified from their bulk values, $p_{B,h}$ ($h = 1,2,3,4$) near a boundary (closed or open). Then one-point functions of operators of dimension d will decay as $1/y^d$, where y is the distance from the boundary, and d is the operator dimension [28, 30].

$$p_h(y) = p_{B,h} + \frac{c_h}{y^2} + \dots, \quad \text{(F1)}$$

for some constants c_h. If the fields are normalized (to have zero expectation value and coeficient -1 in two-point correlations), then general CFT principles predict that the coefficients of the $1/y^2$ terms should be universal numbers, depending only on the field and the boundary condition [28]; in particular, they should be independent of h. So upon normalizing the height variables, the different c_h should all become $\tilde c$, a number independent of h. Since we are assuming that all four height variables are represented by the same field, the 1-1, 2-2, 3-3, and 4-4 correlations should all have the same sign (negative), so this normalization should not change the signs of the coefficients, and all the c_h's should have the same sign as $\tilde c$. However, we need $\sum_{h=1}^{4} c_h = 0$, for the four height probabilites to sum to one, so the c_h cannot all have the same sign. By contradiction, the four height variables must be represented by different fields in the bulk.

[1] P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. Lett., **59**, 381 (1987).

[2] J. M. Carlson and J. S. Langer, Phys. Rev. Lett. **62**, 2632

(1989).

[3] P. Dutta and P.M. Horn, Rev. Mod. Phys. **53**, 497 (1981).

[4] P. S. Dodds, D. H. Rothman, Phys. Rev. E. **59**, 4865 (1999).

[5] P. Bak, *How Nature Works* (Oxford Univ. Press, Oxford, 1997).

[6] H. J. Jensen, *Self-Organized Criticality* (Cambridge University Press, Cambridge, 1998).

[7] D. Dhar, Phys. Rev. Lett. **64**, 1613 (1990); Phys. Rev. Lett. **64**, 2837 (1990).

[8] D. Dhar, Physica A **263**, 4 (1999).

[9] S. N. Majumdar and D. Dhar, Physica A **185**, 129 (1992).

[10] V. B. Priezzhev, J. Stat. Phys. **74**, 955 (1994).

[11] H. Saleur, Nucl. Phys. B **382**, 486 (1992).

[12] V. Gurarie, Nucl. Phys. B **410**, 535 (1993).

[13] M. R. Gaberdiel and H. Kausch, Nucl. Phys. B **538**, 631 (1999).

[14] S. N. Majumdar and D. Dhar, J. Phys. A: Math. Gen. **24**, L357 (1991).

[15] S. Mahieu and P. Ruelle, Phys. Rev. E **64**, 066130 (2001).

[16] E. V. Ivashkevich, J. Phys. A: Math. Gen. **27**, 3643 (1994).

[17] Dhar, "Studying self-organized criticality with exactly solved models," cond-mat/9909009.

[18] M. Jeng, "The four height variables and the Abelian sandpile model," cond-mat/0312656.

[19] J. L. Cardy, Nucl. Phys. B **240**, 514 (1984).

[20] T. Tsuchiya and M. Katori, Phys. Rev. E **61**, 1183 (2000).

[21] P. Ghaffari, S. Lise, and H. J. Jensen, Phys. Rev. E **56**, 6702 (1997).

[22] S. S. Manna, L. B. Kiss, and J. Kertśz, J. Stat. Phys. **61**, 923 (1990).

[23] M. Jeng, Phys. Rev. E. **69**, 051302 (2004).

[24] F. Spitzer, *Principles of Random Walk*, 2nd edition (Springer-Verlag, New York, 1976).

[25] J. G. Brankov, E. V. Ivashkevich, and V. B. Priezzhev, J. Phys. I France **3**, 1729 (1993).

[26] S. N. Majumdar and D. Dhar, J. Phys. A: Math. Gen. **23** 4333 (1990).

[27] G. Piroux and P. Ruelle, "Boundary height fields in the Abelian sandpile model," hep-th/0409126, to be published in Physica A.

[28] J. L. Cardy and D. C. Lewellen, Phys. Lett. B **259**, 274 (1991).

[29] I. I. Kogan and J. F. Wheater, Phys. Lett. B **486**, 353 (2000).

[30] S. Moghimi-Araghi, S. Rouhani, Lett. Math. Phys. **53**, 49 (2000).

[31] S. Kawai and J. F. Wheater, Phys. Lett. B **508**, 203 (2001).

[32] I. I. Kogan and J. F. Wheater, Phys. Lett. B **486**, 353 (2000).

[33] Y. Ishimoto, Nucl. Phys. B **619**, 415 (2001).

[34] A. Bredthauer, M. Flohr, Nucl. Phys. B **639**, 450 (2002).